Characterisation of Keratin-associated Protein Genes Associated with Wool Traits

---

A thesis

submitted in partial fulfilment of the requirements

for

the Degree of Doctor of Philosophy

at

Lincoln University

by

Hua Gong

Lincoln University

New Zealand

2015

Abstract of a thesis submitted in partial fulfilment of
the requirements for the Degree of Doctor of Philosophy

Characterisation of Keratin-associated Protein Genes Associated with Wool Traits

by
Hua Gong

Abstract

In an effort for wool fibre to compete with the synthetics market, the sheep industry is looking for ways to improve their product through selective breeding programmes. The use of gene markers is one way to assist with improving selection of genetically superior stock, with the added benefit of being able to screen animals at an early age without having to wait until they have their adult fleece.

Keratin-associated proteins (KAPs) are key structural components of hair and wool fibres and form a semi-rigid matrix in which the keratin intermediate filaments are embedded. KAPs are believed to play a critical role in determining the physical properties of the fibre and therefore influence key wool traits. Their genes make ideal candidates for the development of gene markers.

In humans, the KAPs are encoded by 88 functional genes, but only a small proportion of KAP genes and variability therein have been identified in sheep.

This study identified variation in five previously identified ovine KAP genes (*KRTAP1-4, KRTAP5-4, KRTAP6-n, KRTAP7-1* and *KRTAP8-1*), revealing nine, five, five, two and five Polymerase Chain Reaction-Single Strand Conformational Polymorphism (PCR-SSCP) banding patterns, respectively.

Either one or a combination of two different patterns was observed for single sheep for all the KAP genes except *KRTAP6-n*. There were between two and five PCR-SSCP patterns observed for *KRTAP6-n* amplicons in each sheep, suggesting the *KRTAP6-n* amplicons were derived from more than one locus.

Nucleotide sequencing of PCR amplicons representative of different PCR-SSCP patterns revealed nine nucleotide sequences for *KRTAP1-4*, five nucleotide sequences for *KRTAP5-4*, five nucleotide sequences for *KRTAP6-n*, two nucleotide sequences for *KRTAP7-1* and five nucleotide sequences for *KRTAP8-1*. There were 14, six, one, and four single nucleotide polymorphisms (SNPs) identified in

KRTAP1-4, KRTAP5-4, KRTAP7-1 and *KRTAP8-1*, respectively. The majority of SNPs were located in the coding region, and nine SNPs in *KRTAP1-4*, four SNPs in *KRTAP5-4*, one SNP in each of *KRTAP7-1* and *KRTAP8-1*, were non-synonymous. Variation in the copy number of a 30 bp sequence encoding a cysteine-rich decapeptide "RPCCSQSSCC", was also identified in *KRTAP5-4*.

Of the five *KRTAP6-n* sequences (*A* to *E*) obtained, sequence *D* was identical to a published ovine *KAP6-1* sequence (GenBank M95719), and sequence *B* was identical to *D* with the exception of a 57 bp deletion/insertion in the coding region and a SNP in the 3'-UTR. These two sequences appear to represent allele variants of ovine *KRTAP6-1*. Sequences *A* and *C* were similar to each other (with only one synonymous SNP), but different to the other sequences. These two sequences appear to be related to a sheep KAP6 amino acid sequence, and represent allelic variation at another *KRTAP6* locus (designated *KRTAP6-2*). The remaining sequence *E* did not show high sequence homology with either the *KAP6-1* or *KAP6-2* sequences, but exhibited homology with a bovine *KAP6-3* sequence, with the exception of a deletion/insertion of 30 nucleotides. This suggests that *E* represents ovine *KRTAP6-3*.

This study also searched for previously unidentified ovine KAP genes and lead to the identification of *KRTAP1-2, KRTAP8-2, KRTAP11-1, KRTAP13-3* and *KRTAP24-1*, all of which were polymorphic. A BLAST search of the Ovine Genome Assembly v.10 using a sequence conserved across the known ovine KAP1-n genes identified a new KAP1-n sequence on chromosome 11. PCR amplification of this sequence revealed an open reading frame of 474 bp that putatively encodes a polypeptide sequence very similar to the previously described ovine KAP1-2 protein. PCR-SSCP analysis revealed nine unique banding patterns, representing nine different nucleotide sequences in *KRTAP1-2*. Ten SNPs were identified and three of the SNPs were non-synonymous.

A BLAST search of the Ovine Genome Assembly v2.0 using the caprine *KRTAP8-2* coding sequence identified a homologous region on sheep chromosome 1. This region was clustered with a number of previously identified KAP genes including (in order from the centromere) *KRTAP11-1*, *KRTAP7-1, KRTAP8-1, KRTAP6-2, KRTAP6-1*, *KRTAP13-3,* and *KRTAP24-1*. PCR-SSCP analysis of this gene revealed two dissimilar PCR-SSCP banding patterns, representing two DNA sequences. These sequences did not have great homology with other known ovine *KRTAP* sequences, but the highest sequence identity was found with *KRTAP8-2* from goats and reindeer, suggesting that the sequences are derived from a newly identified KAP8-2 gene. One SNP located 21 bp upstream of the TATA box was identified. The notional KAP8-2 protein is comprised of 63 amino acid residues, rich in glycine and tyrosine, but has a low cysteine content. In contrast to other High-Glycine-Tyrosine (HGT) KAPs, ovine KAP8-2 contains more acidic amino acid residues, and this would likely result in a lower isoelectric point (pI) of 6.3 which may affect the reaction between amino acid residues.

The putative ovine KAP11-1 gene was identified by PCR using primers designed based on the cattle KAP11-1 gene sequence. Six PCR-SSCP patterns representing six different nucleotide sequences were detected. All of the sequences were unique, and the greatest homology was with *KRTAP11-1* sequences from cattle, human and mouse. This suggested that these sequences were derived from ovine *KRTAP11-1*. The ovine KAP11-1 gene had an open reading frame of 477 nucleotides encoding 159 amino acids. The putative protein was rich in serine, cysteine and threonine, and these account for 18.2 - 18.9, 12.6 and 12.0 mol%, respectively. Of these serine and threonine residues, approximately 20 might be phosphorylated. Five SNPs were identified and one was non-synonymous and would result in an amino acid change at a potential phosphorylation site.

The putative ovine KAP13-3 gene was amplified using primers designed based on a reported bovine *KRTAP13-3* sequence. Five unique PCR-SSCP banding patterns, representing five different nucleotide sequences were identified. Four SNPs were detected among these sequences, and three of them were non-synonymous.

A BLAST search of the Ovine Genome Assembly v2.0 using a human *KRTAP24-1* coding sequence (NM_001085455) identified a putative ovine KAP24-1 gene clustered with other six known KAP genes on chromosome 1. PCR-stem-loop conformational polymorphism (SLCP) analysis of the ovine *KRTAP24-1* revealed four unique banding patterns, representing four different nucleotide sequences. These sequences were not closely homologous to any known ovine *KRTAP* sequence and the highest similarity was with *KRTAP24-1* sequences from other mammalian species, suggesting the sequences were allelic variants of ovine *KRTAP24-1*. Seven SNPs were identified in the coding region and four SNPs were non-synonymous. The putative ovine KAP24-1 polypeptide consisted of 252 amino acids. While probably belonging to the high-sulphur KAP group, the polypeptide had a moderate level of cysteine, but a high content of serine and tyrosine. The polypeptide possesses two putative N-glycosylation sites and a number of residues that may be O-glycosylated and/or phosphorylated.

To accommodate the emerging diversity in KAP genes, an updated KAP nomenclature was proposed. This proposed nomenclature uses the abbreviation sp-KAPm-nL*x for KAP proteins and sp-*KRTAPm-n(p/L)**x for KAP genes. In this system "sp" is a unique letter-based code for different species as described by the protein knowledge-based UniProt; "m" is a number identifying the gene or protein family; "n" is a constituent member of that family; "p" signifies a pseudogene if present; "L" if present signifies "like" and refers to a temporary "place-holder" until the family is confirmed and "x" signifies a genetic variant or allele. The use of non-italicised text for the proteins and italicised text for the genes is supported. This nomenclature is not that different to the existing system, but it includes species information and also describes genetic variation if identified, and hence is more

informative. For example, GenBank sequence JN091630 would historically have been named *KRTAP7-1* for the gene and KAP7-1 for the protein, but with the proposed nomenclature would be SHEEP-*KRTAP7-1*A* and SHEEP-KAP7-1*A for the gene and protein respectively. This nomenclature will facilitate more efficient storage and retrieval of data and define a common language for the KAP proteins and genes from all mammalian species.

The impact of variation in *KRTAP1-2* and *KRTAP6-1* on wool traits was investigated in 383 and 368 Merino × Southdown cross sheep respectively, sired by six different rams. For *KRTAP1-2*, seven previously reported variants (*A-C*, *E-H*) and two newly identified variants (*J* and *K*) were detected, with a frequency of 25.6%, 10.4%, 1.5%, 0.3%, 39.5%, 12.9%, 9.1%, 0.4% and 0.3%, respectively. The effects of variants *E, J* and *K* on wool traits were not investigated as they were found at a very low frequency. Of the other six variants; *A*, *B* and *C* were found to be associated with a number of wool traits, but the strongest association was with three correlated traits: Clean Fleece Weight (CFW), Greasy Fleece Weight (GFW) and yield. The presence of *A* was associated with an increase in GFW, CFW and yield. The presence of *B* was also associated with an increase in GFW, CFW and yield. The presence of *C* was associated with a decrease in GFW and CFW, but increased yield. Genotypes *AF* and *AG* had a higher GFW than *FG* and *FH*; *AF* and *AG* had a higher CFW than *FG*; and sheep of *AF*, *AG* and *BF* had a higher yield than those of *FF* and *FG*.

For *KRTAP6-1*, three gene variants (*A*, *B* and *C*) were identified. Variants *A* and *B* were similar to each other, with only three nucleotide differences occurring downstream of the coding sequence. However, variant *C* had a 57 bp deletion that would notionally result in a loss of 19 amino acids in the protein. The presence of *C* was found to be associated with an increase in Mean Fibre Diameter (MFD), Fibre Diameter Standard Deviation (FDSD), Coefficient of Variation of Fibre Diameter (CVFD) and Prickle Factor (percentage of fibres over 30 microns; PF). Sheep of genotype *BC* produced wool of greater MFD, FDSD and PF than sheep of genotypes *AA*, *AB* and *BB*. The CVFD was greater in the *BC* sheep than the *AB* sheep. The results suggest that variation in ovine *KRTAP6-1* affects wool fibre diameter associated traits and that the 57 bp deletion in this gene would lead to coarser wool with greater FDSD, CVFD and PF.

This study confirms the potential for KAP genes to serve as gene markers of desirable wool traits.

Keywords: Sheep, wool, keratin-associated protein (KAP), keratin-associated protein gene (*KRTAP*), variation, PCR-SSCP, association, fibre diameter, fleece weight.

Publications Arising from this Study

1. Gong H, Zhou H, Hodge S, Dyer JM and Hickford JGH. (2015). Association of wool traits with variation in the ovine KAP1-2 gene in Merino cross lambs. *Small Ruminant Research* 124:24-9.
2. Zhou H*, Gong H*, Li S, Luo Y and Hickford JGH. (2015). A 57-bp deletion in the ovine KAP6-1 gene affects wool fibre diameter. *Journal of Animal Breeding and Genetics* 132:301-7. (*co-first authors).
3. Gong H, Zhou H, Dyer J and Hickford JGH. (2014). The sheep KAP 8-2 gene, a new KAP8 family member that is absent in humans. *SpringerPlus* 3:528.
4. Gong H, Zhou H, McKenzie GW, Yu Z, Clerens S, Dyer JM, Plowman JE, Wright MW, Arora R, Bawden CS, Chen Y, Li J and Hickford JGH. (2012). An updated nomenclature for keratin-associated proteins (KAPs). *International Journal of Biological Sciences* 8:258-64.
5. Zhou H, Gong H, Yan W, Luo Y and Hickford JG. (2012). Identification and sequence analysis of the keratin-associated protein 24-1 (KAP24-1) gene homologue in sheep. *Gene* 511:62-5.
6. Gong H, Zhou H, Dyer JM, Plowman JE and Hickford JGH. (2012). Identification of the keratin-associated protein 13-3 (KAP13-3) gene in sheep. *Open Journal of Genetics* 1:60-4.
7. Gong H, Zhou H, Plowman JE, Dyer JM and Hickford JGH. (2012). Search for variation in the ovine KAP7-1 and KAP8-1 genes using polymerase chain reaction-single-stranded conformational polymorphism screening. *DNA and Cell Biology* 31:367-70.
8. Gong H, Zhou H, Yu Z, Dyer J, Plowman JE and Hickford JGH. (2011). Identification of the ovine keratin-associated protein KAP1-2 gene (*KRTAP1-2*). *Experimental Dermatology* 20: 815-9.
9. Gong H, Zhou H, Dyer JM and Hickford JGH. (2011). Identification of the ovine KAP11-1 gene (*KRTAP11-1*) and genetic variation in its coding sequence. *Molecular Biology Reports* 38: 5429-33.
10. Gong H, Zhou H and Hickford JGH. (2011). Diversity of the glycine/tyrosine-rich keratin-associated protein 6 gene (*KAP6*) family in sheep. *Molecular Biology Reports* 38:31-5.
11. Gong H, Zhou H, McKenzie GW, Hickford JGH, Yu Z, Clerens S, Dyer JM and Plowman JE. (2010). Emerging issues with the current keratin-associated protein nomenclature. *International Journal of Trichology* 2:104-5.
12. Gong H, Zhou H, Plowman J, Dyer J and Hickford JGH. (2010). Analysis of variation in the ovine ultra-high sulphur keratin-associated protein KAP5-4 gene using PCR-SSCP technique. *Electrophoresis* 31:3545-7.
13. Gong H, Zhou H and Hickford JGH. (2010). Polymorphism of the ovine keratin-associated protein 1-4 gene (*KRTAP1-4*). *Molecular Biology Reports* 37:3377-80.

Nucleotide Sequences Submitted to the NCBI GenBank

SHEEP-*KRTAP1-2*: HQ897973-HQ897982, KM105941-KM105942

SHEEP-*KRTAP1-4*: GQ507741-GQ507749

SHEEP-*KRTAP5-4*: GU255997-GU256001

SHEEP-*KRTAP6-n*: GU319872-GU319876

SHEEP-*KRTAP7-1*: JN091630-JN091631

SHEEP-*KRTAP8-1*: JN091632-JN091636

SHEEP-*KRTAP8-2*: KF220646-KF220647

SHEEP-*KRTAP11-1*: HQ595347-HQ595352

SHEEP-*KRTAP13-3*: JN377429-JN377433

SHEEP-*KRTAP24-1*: JX112014- JX112017

Acknowledgements

The completion of this PhD thesis was only possible with the support and assistance received from many generous and kind people. I thank you all here. My apologies if I miss anyone out.

Firstly, I would like to thank my main supervisor, Professor Jon Hickford for his continuous support and supervision during this research. A huge thanks not only for your acedemic support, but also for the continous motivation and encouragement of me. Thank you very very much, Jon!

I am indebted to my Co-supervisor, Associate Professor Jolon Dyer, for his continuous support during the duration of my Ph.D. studies. I am especially grateful for your reference letter which helped me secure my three-year scholarship.

No PhD is possible without a suitable source of funding. The Wool Research Organisation of New Zealand Inc and the New Zealand Wool Industry Charitable Trust provided me with a three-year Postgraduate Scholarship. I would also like to acknowledge the assistance of funding received from the Foundation for Research, Science and Technology (C10X0710: Keeping New Zealand Wool Products at the Cutting Edge through Enhanced Wool Quality) and the Lincoln University Gene-Marker Laboratory.

Many people helped me along the way. I'd like to thank Freeman Fang and Huitong Zhou for their advice on experimental work, such as optimising PCR and SSCP conditions. Drs Jeffery E. Plowman and Grant McKenzie, your critical input into my writing skills and the interpretation of my results was invaluable. Andrea Hogan, your contribution to the wool database provided valuable information in helping my data analysis. Thanks to Dr Simon Hodge for his help with the statistical analyses, and to Colin Pettigrew and other farm members at Ashley Dene for taking great care of the flock of sheep that was used in the project. Many thanks to the shearers and everyone else who helped out at shearing: Seung-Ok Byun, Guo Yang, Wei Yan, Dr Huitong Zhou and Professor Jon Hickford. Thanks also to John Bates, a sheep industry consultant.

I would like to acknowledge and thank the New Zealand Wool Testing Authority Ltd (NZWTA) for the measurements of wool traits required in this research.

I would like to thank members of the Gene Marker Laboratory (Freeman Fang, Andrea Hogan, Wenqiong Cai, Qingming An, Yunhai Li, and Drs Seung-Ok Byun, Huitong Zhou, Jiqing Wang, Wei Yan, Guo Yang and Jinzhong Tao) and the staff and postgraduates of the Faculty of Agriculture and Life Sciences who have helped me in many ways during the duration of my PhD study, shown me great kindness and contributed positively to my time at Lincoln.

I would also like to thank Dr Lucette Hogg for her wonderful editing and proof reading of this thesis.

I would like to thank my friends in both New Zealand and China, who have supported me during the duration of my PhD. Special thanks to Mimi and Patrick Van Gestel, thank you so much for your friendship and generosity. A huge thanks to Shuxia Hong, Winnie Limcheney, Judie Zhu, Charles Yu, En Hui Wu-li and Chen Wu. Without your continous support, it would have been impossible to carry on with my PhD study.

I am very thankful to my parents for their constant love and support. Thank you, Mum and Dad, for an upbringing that installed in me the concept of hard work, perseverance, and determination in life with regard to achieving set goals. To both of my sisters, I thank you for your support and wise advice that kept me sane during my toughest moments. Your regular communication with me really cheered me up. I love you all.

Finally, I thank my husband for the support and sacrifices you have made over the past years during my studies. To my lovely daughters, Nadia, I thank you especially for your encouragement and positive thinking. Ellie, I thank you for the happiness and surprises you brought me along the way. I love you all.

Table of Contents

List of Tables

List of Figures

Abbreviations

%	percent
°C	degree celsius
°/mm	degree per millimeter
α	alpha
β	beta
σ	sigma
π	pi
µg	microgram
µL	microlitre
µm	micrometre
µM	micromolar
A	adenine
ATLAS	Automatic Tester for Length and Strength
ATP	Adenosine triphosphate
AWI	Australian Wool Innovation Ltd
AWTA	Australia Wool Testing Authority
bp	base pair
C	cytosine
CF	comfort factor
CFW	clean fleece weight
cM	centimorgan
CVFD	coefficient of variation of fibre diameter
Da	dalton
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
DMF	dark and medullated fibre
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
dTTP	deoxythymidine triphosphate
eBV	estimated breeding value
EDTA	ethylenediaminetetraacetic acid
EST	expressed sequence tag
FAOSTAT	Food and Agricultural Organization of the Union Nations Statistics Division
FDSD	fibre diameter standard deviation
g	gram
g	gravity
G	guanine
GFW	greasy fleece weight
GLM	general linear model
h	hour
HGNC	Human Genome Nomenclature Committee
HGT	high glycine-tyrosine
HGVS	human genome variation society
HS	high sulphur
HWE	Hardy-Weinberg equilibrium
IWTO	International Wool Textile Organization
IF	intermediate filament
IRS	inner root sheath
KAP	keratin associated protein
K	keratin
kb	kilobase
kDa	kilodalton
KIF	keratin intermediate filament
KRT	keratin gene

KRTAP	keratin associated protein gene
M	molar
Mb	megabase
MFC	mean fibre curvature
MFD	mean fibre diameter
mg	milligram
micron	micrometer
min	minute
mL	millilitre
mm	millimetre
mM	millimolar
mRNA	messenger RNA
MSL	mean staple length
MSS	mean staple strength
NCBI	national centre for biology information
N/ktex	newtons per kilotex
ng	nanogram
nm	nanometre
nt	nucleotide
NZWTA	New Zealand wool testing authority
OFDA	optical fibre diameter analysis
ORS	outer root sheath
P	primary
PCR	polymerase chain reaction
PF	prickle factor
pI	isoelectric point
QTL	quantitative trait loci
RFLP	restriction fragment length polymorphism
r_g	genetic correlation
r_e	environmental correlation
RNA	ribonucleic acid
r_p	phenotypic correlation
rRNA	ribosomal RNA
S^o	secondary original
S^d	secondary derived
S	second
SCMK	S-carboxy methyl kerateine
SE	standard error
SLCP	stem loop conformational polymorphism
SNP	Single nucleotide polymorphism
SSCP	single strand conformational polymorphism
T	thymine
Taq	*Thermus aquaticus*
TBE	tris-borate-EDTA
TE	tris-EDTA
Tris	tris(hydroxylethyl)aminomethane
U	unit
UHS	ultra high sulphur
UTR	untranslated region
UV	ultraviolet
V	volt
VM	vegetable matter

Amino Acid Residue Abbreviations

A	**Ala**	alanine
C	**Cys**	cysteine
D	**Asp**	aspartic acid
E	**Glu**	glutamic acid
F	**Phe**	phenylalanine
G	**Gly**	glycine
H	**His**	histidine
I	**Ile**	isoleucine
K	**Lys**	lysine
L	**Leu**	leucine
M	**Met**	methionine
N	**Asn**	asparagine
P	**Pro**	proline
Q	**Gln**	glutamine
R	**Arg**	arginine
S	**Ser**	serine
T	**Thr**	threonine
V	**Val**	valine
W	**Trp**	tryptophan
Y	**Tyr**	tyrosine

Chapter 1

Literature Review

1.1 Introduction

Wool, a natural fibre with some very unique attributes, is widely used in the textile, insulation and carpet industries. Humans have long recognised the benefits of wool including its thermal attributes, breathability and fire resistance (Leeder, 1984). Wool, however, is facing growing competition from other textile fibres, especially synthetic fibres. Synthetic fibres can be engineered to have specific properties and their production can be easily tailored to demand. Furthermore, scales of economic in the production of man-made fibres provide for an efficient cost-effective manufacturing process.

In comparison, the wool industry's reliance on a natural product which is subject to all kinds of variability, has more difficulty delivering stable production volumes and uniform fibre quality. Wool fibres vary in quality among sheep breeds, between sheep of one breed, within single fleeces and individual staples and fibres. Wool production also takes time and requires a consistently healthy and nutritionally adequate environment. For the best results. Possibly as a results of these limitations, during 1994 - 2004, demand for wool declined by nearly 30% in terms of its share of the total volume of global textile fibre consumption (Cottle, 2010).

To meet the competition posed by synthetics, the wool industry has to take on the task of improving wool production and quality through selective breeding programmes. This study contributes to this objective by identifying potential genetic markers for beneficial wool characteristics allowing the rapid identification of genetically superior animals.

1.1.1 Wool

Wool describes the fine curly hair that constitutes the fleece produced by domestic sheep (*Ovis aries*). Less commonly, it is also used as the generic name of hair from other animals such as goat, alpaca, yak, camel, lama and rabbit (Hocker, 2002).

1.1.2 Unique Attributes of Wool

Wool has outstanding insulation properties. The wool fibres are crimped, giving wool a three dimensional structure that inhibits the fibres from packing together. As a result, small pockets of air are trapped between the fibres. This air trapped inside the fabric acts as an excellent insulator. This means that in cold weather it keeps the body warm, whereas in hot weather it keeps the body cool.

Wool has a hydrophobic exterior and hydrophilic interior which confer unique moisture absorption properties (Simpson & Crawshaw, 2002). For sheep this prevents moisture from being held at the skin surface. In woollen clothing it prevents the clammy, cold feeling you may experience when wearing some types of synthetic clothing. Because of this feature, wool also dries out relatively slowly (Leeder, 1984).

Wool is relatively free of static charge build-up because its chemical structure and water-absorbing properties make it a good conductor (Wilkinson, 1970). It also has relatively good flame resistance and is difficult to ignite (Leeder, 1984). Any flame produced spreads slowly and is easily extinguished. The residue left after burning is a low-temperature fragile non-sticking ash. Synthetic fibres on the other hand, frequently burn hotly and can produce hot molten residue (Leeder, 1984). These features make wool attractive as an insulating material in homes and other interior textile applications, as well as in aviation applications.

As a consequence of having crimp, wool has the ability to stretch and then return to its original state. This elasticity is a promising feature that is difficult to achieve with any other synthetic material (Leeder, 1984). This makes the fibre both strong and durable under repeated loading.

Another feature making wool strong and durable is its ability to felt. The felting process ensures the fibres hold together during twisting and turning of the textile. This feature assists in allowing wool to be used in creating long lasting garments (Leeder, 1984).

Wool fibres readily accept dyes. The natural off-white raw colour is easily changed through dyeing because of the wide variety of chemical groups available for dyes to bind to the fibre (Leeder, 1984).

1.1.3 Worldwide Production of Wool

There are more than 200 sheep breeds in the world and the total sheep population was over 1.1 billion animals in 2013 (FAOSTAT, 2015). Domestic sheep are bred for many purposes including the production of meat, milk, leather and wool. In 2013, world greasy wool production was about 2.12 million tonnes per annum (FAOSTAT, 2015). While China, Australia and India have the highest sheep numbers, Australia, China and New Zealand are the top three wool producing countries (FAOSTAT, 2015). In 2013, China produced about 22% of the total world greasy wool production (360,520 tons), while Australia and New Zealand produced nearly 25% of the world greasy wool production (525,520 tons) (FAOSTAT, 2015).

World wool production is declining in spite of an increase in the world sheep population (FAOSTAT, 2015). This is due to a shift towards dual-purpose (meat and wool) sheep breeds at the expense of

wool breeds. It means that the supply of wool is very tight and the lowest in at least 70 years (FAOSTAT, 2015). China has been taking various steps to boost its economy and the demand for wool is growing. It has become the largest importer of raw wool (39% of the world total imports in 2006) and the largest producer of wool textiles in the world (over 50% of the world total) (IWTO, 2015).

As a textile fibre, about 60% of world wool production is processed into apparel, the main destination of fine wool (IWTO, 2015). The other products for which wool is used are interior textiles and technical textiles. The suitability of wool for any given purposed is determined by specific wool traits.

1.2 Wool Traits and Their Measurement

Wool traits are the characteristics of wool fibres affecting physical processing performance and the quality of the product produced. Wool is not a uniform biological product and its traits can vary dramatically depending on sheep genetics, environment and flock management.

Typically for the buyer, fibre diameter is recognised as the most important trait in determining wool price and value (AWI, n.d.). Other economically important wool quality traits include staple length, staple strength, colour, curvature, softness of handle, dark and medullated fibre (DMF) contamination, and vegetable matter (VM) content (AWI, n.d.). For the farmer, the amount of wool produced per sheep obviously also affects profitability (as well as the quality of the fibre).

Fibre Diameter

Because fibre diameter is such a driving determinant of wool quality, wool is classified into three basic groups: fine wool, medium wool and coarse or strong wool. Merino is the main breed producing high quality fine wool in New Zealand and Australia. According to the International Wool Textile Organisation (IWTO), 37% of world wool production is classified as fine wools, 22% of world wool production is classified as medium wools, and the remainder at 41% is classified as coarse wool (IWTO, 2015).

It is not common to measure the diameter of individual fibres as typically wool grows in clusters of fibres called staples, so typically the Mean Fibre Diameter (MFD) is measured. This refers to the average width of a fibre cross section and is expressed in micrometres (microns or µm). Based on MFD measurement, wool fibre can be categorised as ultra-fine wool (less than 16 µm), super-fine wool (16 µm to 19 µm), fine wool (19 µm to 23 µm), medium wool (23 µm to 32 µm), or coarse wool (over 32 µm).

MFD is widely acknowledged as the most influential factor on both price and end use of wool. It accounts for about 75% of the total price of raw wool (Cottle, 2010) and price premium typically increases with wool fineness. MFD affects processing performance and the quality of the final products made with wool. As a minimum number of fibres must be present in the cross-section of any commercially acceptable yarn, if these fibres are finer, lighter yarns may be spun making for softer and lighter fabrics. Such fabrics usually have superior handle and drape in garments. Yarn evenness is also affected by fibre diameter, particularly in the worsted spinning process (Cottle, 1991). Thus fine wools with low MFD are suited for high value apparel textile end uses. Coarser wools with higher MFD values are better suited for less luxurious and lower valued uses such as making carpet, outerwear or bedding.

MFD has been measured by the Airflow method (IWTO-28) (IWTO, 2013) for many years. This method works on a pressure drop principle. In the airflow instrument, air is forced through the fibre sample. The greater the pressure drop, the finer the wool. However, the Airflow method cannot provide information about variation in diameter along and between individual fibres. Two technologies, LASERSCAN and Optical Fibre Diameter Analysis (OFDA), have therefore been developed to provide this additional information and they have replaced the Airflow method for many applications (Botha & Hunter, 2010). LASERCAN measures the change in the light generated by a laser when the fibre snippets pass through a measurement cell, and the amount of light incident is directly proportional to the projected area and therefore thickness or diameter of the fibre. OFDA is essentially an automatic microscope set above a prepared slide of fibre snippets, and it can rapidly determine fibre diameter distribution.

As OFDA or LASERSCAN assess a population of fibre snippets, fibre diameter values from either are best represented as distributions, allowing the evaluation of the degree of variation in the sample (Aylan-Parker & McGregor, 2002; Botha & Hunter, 2010). Variation can be expressed in terms of Fibre Diameter Standard Deviation (FDSD) or Coefficient of Variation of Fibre Diameter (CVFD). CVFD is the FDSD divided by MFD multiplied by 100%.

Comfort Factor (CF; percentage of fibre of less than or equal to 30µm in diameter) or Prickle Factor (PF; the opposite of CF) can also be calculated from these fibre snippet populations since these traits are largely dependent on fibre diameter. Evidence suggests that protruding wool fibres with diameters less than 30 µm are deflected upon contact with the skin and avoid irritation (Naylor, 2010). Usually PF is kept below 5% to improve the product value and marketability (Malau-Aduli & Deng Akuoch, 2010; Naylor *et al.*, 1995). Premiums are typically offered for wools with low FDSD, low CVFD, high CF and low PF.

Fibre diameter also has an influence on dyeing performance. Finer fibres yield darker shades in both yarn and product form because of a higher surface area to volume ratio than coarser fibres.

Mean Staple Length and Mean Staple Strength

Mean Staple Length (MSL) and Mean Staple Strength (MSS) are another two wool traits imparting wool quality. The importance of MSL and MSS as wool traits is linked to their positive relationship with wool processing performance. MSS is especially recognised as second only to MFD in determining the value of fine wool (Botha and Hunter, 2010; Brown *et al.*, 2002). Wools with higher MSL and MSS are more commercially desirable as they tend to be easier to spin, result in few machine stoppages and ultimately can form stronger and more even yarns (Angel *et al.*, 1990; Cottle, 1991). Especially in the early stages of wool processing, lower MSS wools can produce shorter fibres, just like lower MSL wools. This typically results in surface fuzzing and pilling in apparel fabric surfaces, and fibre loss from woollen carpets (Cottle, 1991).

Although higher MSL and MSS are desirable wool traits, there are two wool processing systems to deal with wool of variable MSL and MSS: the woollen and worsted systems. The longer wools, generally around 65 mm or longer are called combing types and these are processed to worsted yarn (Cottle, 1991). The woollen system handles wool of any length and strength, as it produces bulky yarns of low twist levels to produce a wide range of wool products. In contrast, the worsted system requires wool of higher length and strength. It produces smooth compact yarns of reasonably uniform length and fineness. These are used to produce light or medium weight woven and knitted fabrics for luxury garment manufacture.

MSL and MSS are measured by The Automatic Tester for Length and Strength (ATLAS) developed by the Australia Wool Testing Authority Ltd (AWTA). This instrument has been used in New Zealand since 1990 (Wools of New Zealand, 1996).

MSL is measured in units of millimetres (mm) or inches without stretching the wool staple. MSS is expressed as the force required to break a staple of a given thickness and measured in Newtons/Kilotex (N/ktex). It varies from less than 10 N/ktex to over 80 N/ktex (Cottle, 2010). Commercially, high MSS wools are called “sound wools” and are considered of markedly greater value than their low MSS or “tender” wool counterparts (Cottle, 2010). Wools are therefore routinely allotted into one of four main descriptive categories: 1) Sound, which includes staples stronger than 25 - 30 N/ktex; 2) Part-tender, with MSS approximating 20 N/ktex; 3) Tender, with MSS values around 15 N/ktex; and 4) Rotten, referring to staples breaking with less than 10 N/ktex of force (Cottle, 2010).

Fibre Curvature/Crimp

Crimp, a characteristic of wool that distinguishes it from hair and fur, refers to the natural curl in the wool fibre. Over recent decades, wool fibre crimp has been described in terms of fibre curvature (rotation of the fibre around a central axis. The higher the crimp, the higher the degree of curvature. Wool fibre curvature is expressed in degrees per millimetre (°/mm) fibre length. Mean Fibre Curvature (MFC) is the average curvature of fibre snippet measured by the OFDA or LASERSCAN. Mean curvature values range from over 100 °/mm for fine merinos, to less than 40 °/mm for coarse cross-bred wool.

In wool processing, higher curvature is less desirable because it may impact on the probability of fibre entanglement, resulting in more frequent fibre breakage. This in turn, causes more wool wastage and produces more protruding fibres similar to the low MSS and MSL wool. However, wool with high curvature can promote fibre cohesion in yarn thereby assisting in processing performance. It also gives the fibre elasticity and resilience that enhance the thermal insulation qualities of the end products. Overall, optimum wool curvature depends on the processing technology and the end product application.

Colour

Colour of most greasy wool varies from white through to shades of cream, yellow and brown. Wool colour is usually measured by colorimetry using a reflectance spectrophotometer to obtain tristimulus values for light reflected from wool at three different wavelengths of the spectrum referred to as X, Y and Z. The X value represents red colour, the Y value represents green colour and the Z value represents blue colour. The value of X or Y describes brightness (the higher the number the brighter the sample). Since yellow colour is the perception of the difference between green and blue, the difference between Y and Z describes the mean yellowness (the higher the number, the more yellow the wool). In general, wool brightness values vary from 56 to 66 and the yellowness values range from -5 to +15. Wool colour is somewhat related to the fibre diameter. Generally fine wools are whiter with the colour becoming creamier as the fibre diameter increases.

Bright and white wool colour is usually preferred commercially because this kind of wool can be dyed consistently to a large range of shades (Reid & Botica, 1995). Yellow and dull wool has limitations in dyeing processes. Some colour in greasy wool can be removed by scouring, but this increases the cost of processing. Non-scourable colour and dullness in clean wool precludes wool from being dyed to light or pastel shades and typically coloured wools are discounted in the market.

Fleece Weight and Yield

Greasy Fleece Weight (GFW) is the mass of raw wool just shorn from the sheep. After removing the impurities such as vegetable matter, wax, suint, dirt and dung by scouring, the mass becomes the Clean Fleece Weight (CFW). The Yield is expressed as the CFW divided by GFW as a percentage. Lower yield is less desirable as it will increase the cost of wool processing.

The GFW is typically measured after shearing. The wool grower can be paid on the basis of this indirect criterion as GFW and CFW are positively correlated (Table 1-1).

1.2.1 The Heritability of Wool Traits

Heritability (h^2) is an indication of the extent that variance in progeny for a given trait can be explained by variance in the parental population for that trait. It is expressed on a scale of 0.0 to 1.0. A trait with a heritability value of 0.3 or more is regarded as moderate to highly heritable. The use of such parents in breeding programmes will lead to improvements in the trait of the offspring, providing the animals have been reared together to eliminate environmental impacts on expression of this trait (Turner, 1964).

Previous research indicates that wool traits have moderate to high heritability ranging from 0.3 to 0.6 (Table 1-1). This suggests that wool traits are under genetic control and that flocks can be improved by selective breeding. This is not without precedence as it is understood that sheep have been bred for wool production since 8000 BC.

1.2.2 Correlations between Various Wool Traits

In animal breeding, the improvement of one trait may have a positive or negative effect on other traits. There are two kinds of correlations measured, phenotypic and genetic correlations. Phenotypic correlations (r_p) measure different characteristics on the same animal, while genetic correlations (r_g) measure characteristics shared by an animal and its relatives. Phenotypic correlations are observed between traits and depend on both environmental correlations (r_e) and genetic correlations. The relationship between these three is not simple and they may differ in sign, as well as in value, for the same pair of traits (Bowman, 1974).

A genetic correlation can be the result of pleiotropy whereby a single gene is influencing more than one trait at the same time, or because of linkage between two or more gene loci, each of which is influencing a single character (Johansson & Rendel, 1968). A genetic correlation caused by pleiotropy is permanent and cannot be altered by selection, whereas one caused by linkage is transient and may change in time as a consequence of selection and crossing-over (recombination) during meiosis (Bowman, 1974).

Table 1-1. Correlation estimates reported for wool traits.

Traits		Breed*	Genetic correlation (± SE)$^{\Phi}$	Phenotypic correlation (± SE)$^{\Phi}$	Reference
GFW ×	**CFW**	Merino	0.84 ± 0.05	0.92 ± 0.01	(Wuliji *et al.*, 2001)
		Merino	0.89 ± 0.01	0.79 ± 0.01	(Safari *et al.*, 2007)
		Romney	0.95 ± 0.01	0.95 ± 0.00	(Wuliji *et al.*, 2011)
		Corriedale	0.89 ± 0.05	0.92 ± 0.01	(Benavides & Maher, 2000)
	Yield	Merino	-0.18 ± 0.02	-0.07 ± 0.01	(Safari *et al.*, 2007)
		Romney	-0.03 ± 0.10	0.01 ± 0.02	(Wuliji *et al.*, 2011)
		Corriedale	-0.29 ± 0.22	-0.06 ± 0.06	(Benavides & Maher, 2000)
	MFD	Merino	0.27 ± 0.02	0.24 ± 0.01	(Safari *et al.*, 2007)
		Romney	0.40 ± 0.07	0.48 ± 0.02	(Wuliji *et al.*, 2011)
		Corriedale	0.58 ± 0.16	0.46 ± 0.05	(Benavides & Maher, 2000)
	FDSD	Merino	0.22 ± 0.02	0.17 ± 0.01	(Safari *et al.*, 2007)
	CVFD	Merino	0.10 ± 0.02	0.05 ± 0.01	(Safari *et al.*, 2007)
	MSS	Merino	0.31 ± 0.11	0.25 ± 0.03	(Wuliji *et al.*, 2001)
		Romney	0.31 ± 0.11	0.25 ± 0.03	(Wuliji *et al.*, 2011)
	MSL	Romney	0.44 ± 0.09	0.38 ± 0.02	(Wuliji *et al.*, 2011)
	Brightness	Romney	0.11 ± 0.15	0.01 ± 0.02	(Wuliji *et al.*, 2011)
		Corriedale	-0.07 ± 0.33	-0.05 ± 0.06	(Benavides & Maher, 2000)
	Yellowness	Romney	0.08 ± 0.14	0.01 ± 0.02	(Wuliji *et al.*, 2011)
		Corriedale	0.44 ± 0.31	0.13 ± 0.05	(Benavides & Maher, 2000)
CFW ×	**Yield**	Merino	0.28 ± 0.02	0.35 ± 0.01	(Safari *et al.*, 2007)
		Merino	0.31 ± 0.03	0.54 ± 0.11	(Wuliji *et al.*, 2001)
		Romney	0.23 ± 0.09	0.30 ± 0.02	(Wuliji *et al.*, 2011)
		Corriedale	0.18 ± 0.24	0.34 ± 0.05	(Benavides & Maher, 2000)
	MFD	Merino	0.29 ± 0.02	0.24 ± 0.01	(Safari *et al.*, 2007)
		Merino	0.15 ± 0.12	0.21 ± 0.03	(Wuliji *et al.*, 2001)
		Romney	0.42 ± 0.07	0.50 ± 0.02	(Wuliji *et al.*, 2011)
		Corriedale	0.79 ± 0.14	0.50 ± 0.04	(Benavides & Maher, 2000)
	FDSD	Merino	0.16 ± 0.03	0.13 ± 0.01	(Safari *et al.*, 2007)
	CVFD	Merino	0.01 ± 0.02	0.01 ± 0.01	(Safari *et al.*, 2007)
	MSS	Merino	0.36 ± 0.28	0.15 ± 0.04	(Wuliji *et al.*, 2001)
		Romney	0.41 ± 0.11	0.29 ± 0.03	(Wuliji *et al.*, 2011)
	MSL	Merino	0.21 ± 0.14	0.30 ± 0.04	(Wuliji *et al.*, 2001)
		Romney	0.55 ± 0.08	0.42 ± 0.02	(Wuliji *et al.*, 2011)
	Brightness	Merino	0.21 ± 0.14	0.10 ± 0.03	(Wuliji *et al.*, 2001)
		Romney	0.18 ± 0.15	0.03 ± 0.02	(Wuliji *et al.*, 2011)
		Corriedale	-0.06 ± 0.36	-0.02 ± 0.05	(Benavides & Maher, 2000)
	Yellowness	Merino	0.24 ± 0.14	0.08 ± 0.03	(Wuliji *et al.*, 2001)
		Romney	0.06 ± 0.14	0.10 ± 0.02	(Wuliji *et al.*, 2011)

		Corriedale	0.91 ± 0.24	0.16 ± 0.05	(Benavides & Maher, 2000)
Yield ×	**MFD**	Merino	0.06 ± 0.02	-0.00 ± 0.01	(Safari *et al.*, 2007)
		Merino	0.09 ± 0.09	0.00 ± 0.03	(Wuliji *et al.*, 2001)
		Romney	0.17 ± 0.08	0.13 ± 0.03	(Wuliji *et al.*, 2011)
		Corriedale	0.23 ± 0.19	0.17 ± 0.06	(Benavides & Maher, 2000)
	FDSD	Merino	-0.08 ± 0.03	-0.07 ± 0.01	(Safari *et al.*, 2007)
	CVFD	Merino	-0.14 ± 0.03	-0.10 ± 0.01	(Safari *et al.*, 2007)
	MSS	Merino	0.52 ± 0.22	0.13 ± 0.04	(Wuliji *et al.*, 2001)
		Romney	0.32 ± 0.11	0.10 ± 0.03	(Wuliji *et al.*, 2011)
	MSL	Merino	0.19 ± 0.12	0.22 ± 0.04	(Wuliji *et al.*, 2001)
		Romney	0.30 ± 0.10	0.12 ± 0.03	(Wuliji *et al.*, 2011)
	Brightness	Merino	0.29 ± 0.11	0.13 ± 0.03	(Wuliji *et al.*, 2001)
		Romney	0.19 ± 0.14	0.06 ± 0.02	(Wuliji *et al.*, 2011)
		Corriedale	0.01 ± 0.28	0.06 ± 0.06	(Benavides & Maher, 2000)
	Yellowness	Merino	0.35 ± 0.11	0.12 ± 0.03	(Wuliji *et al.*, 2001)
		Romney	-0.21 ± 0.14	-0.05 ± 0.02	(Wuliji *et al.*, 2011)
		Corriedale	0.36 ± 0.25	0.09 ± 0.06	(Benavides & Maher, 2000)
MFD ×	**FDSD**	Merino	0.54 ± 0.01	0.41 ± 0.01	(Safari *et al.*, 2007)
		NZ-DP	0.78 ± 0.12	0.64 ± 0.02	(Pickering *et al.*, 2013)
	CVFD	Merino	-0.16 ± 0.02	-0.10 ± 0.01	(Safari *et al.*, 2007)
		NZ-DP	0.23 ± 0.24	0.10 ± 0.03	(Pickering *et al.*, 2013)
	MSS	Merino	0.52 ± 0.25	0.25 ± 0.04	(Wuliji *et al.*, 2001)
		Romney	0.52 ± 0.09	0.32 ± 0.03	(Wuliji *et al.*, 2011)
	MSL	Merino	0.00 ± 0.12	0.07 ± 0.05	(Wuliji *et al.*, 2001)
		Romney	0.34 ± 0.09	0.35 ± 0.03	(Wuliji *et al.*, 2011)
	Curvature	NZ-DP	-0.57 ± 0.16	-0.55 ± 0.02	(Pickering *et al.*, 2013)
	Brightness	Merino	0.16 ± 0.11	0.04 ± 0.03	(Wuliji *et al.*, 2001)
		Romney	-0.19 ± 0.14	-0.03 ± 0.02	(Wuliji *et al.*, 2011)
		Corriedale	-0.42 ± 0.25	-0.18 ± 0.05	(Benavides & Maher, 2000)
	Yellowness	Merino	0.62 ± 0.09	0.37 ± 0.03	(Wuliji *et al.*, 2001)
		Romney	0.45 ± 0.11	0.15 ± 0.02	(Wuliji *et al.*, 2011)
		Corriedale	0.26 ± 0.05	0.93 ± 0.18	(Benavides & Maher, 2000)
FDSD ×	**CVFD**	Merino	0.81 ± 0.01	0.79 ± 0.01	(Safari *et al.*, 2007)
		NZ-DP	0.83 ± 0.10	0.82 ± 0.01	(Pickering *et al.*, 2013)
	MSL	NZ-DP	0.24 ± 0.15	0.21 ± 0.03	(Pickering *et al.*, 2013)
	Curvature	NZ-DP	-0.31 ± 0.25	-0.31 ± 0.03	(Pickering *et al.*, 2013)
CVFD ×	**MSL**	NZ-DP	0.11 ± 0.15	0.08 ± 0.03	(Pickering *et al.*, 2013)
	Curvature	NZ-DP	0.03 ± 0.28	-0.03 ± 0.03	(Pickering *et al.*, 2013)
MSS ×	**MSL**	Merino	0.35 ± 0.30	0.03 ± 0.05	(Wuliji *et al.*, 2001)
		Romney	0.46 ± 0.11	0.34 ± 0.02	(Wuliji *et al.*, 2011)
MSL ×	**Curvature**	NZ-DP	-0.71 ± 0.10	-0.37 ± 0.03	(Pickering *et al.*, 2013)

Brightness ×Yellowness	Merino	-0.16 ± 0.13	-0.14 ± 0.03	(Wuliji *et al.*, 2001)
	Romney	-0.62 ± 0.14	-0.38 ± 0.02	(Wuliji *et al.*, 2011)
	Corriedale	-0.64 ± 0.22	-0.50 ± 0.04	(Benavides & Maher, 2000)

* NZ-DP: New Zealand dual-purpose sheep which are dominated by Romney, Coopworth, Perendale, Texel and composite crosses of these breeds.

$^{\phi}$ SE: standard error.

The correlation coefficient, denoted by r, measures the strength and direction of a linear relationship between two traits, and the value of r may vary from -1.0 to +1.0, with the former end indicating the highest negative relationship and the latter end the highest positive one. Correlations with $|r| > 0.7$ are regarded as strong, those with $0.3 < |r| \leq 0.7$ are regarded as moderate and those with $|r| \leq 0.3$ are regarded as weak (Ratner, n.d.). Positive and negative relationships can be either desirable or undesirable.

There are many genetic and phenotypic correlations between wool traits, as shown in Table 1-2. Overall, strong positive correlations are observed between GFW and CFW, and between FDSD and CVFD. Moderate correlations are observed among GFW, CFW and MFD, and among GFW, CFW, MSS and MSL.

Understanding genetic correlations allows prediction of the change in a characteristic that is not directly selected for when selecting for a correlated trait, thereby ensuring the overall effect of breeding is enhanced instead of being offset. This understanding is an integral part in developing better breeding strategies for improving wool quality and ultimately farm profitability.

Table 1-2. Heritability (h^2) estimates for wool traits.

Trait	Breed*	Heritability (± SE)$^{\phi}$	Reference
GFW	Merino	0.24 ± 0.07	(Wuliji *et al.*, 2001)
	Merino	0.29 ± 0.06	(Mortimer & Atkins, 1989)
	Romney	0.44 ± 0.16	(Hawker *et al.*, 1988)
	Romney	0.35 ± 0.04	(Wuliji *et al.*, 2011)
	Corriedale	0.52 ± 0.15	(Benavides & Maher, 2000)
CFW	Merino	0.54 ± 0.10	(Gifford *et al.*, 1995)
	Merino	0.28 ± 0.07	(Wuliji *et al.*, 2001)
	Merino	0.30 ± 0.06	(Mortimer & Atkins, 1989)
	Romney	0.46 ±0.16	(Hawker *et al.*, 1988)
	Romney	0.36 ± 0.04	(Wuliji *et al.*, 2011)

	Corriedale	0.37 ± 0.15	(Benavides & Maher, 2000)
Yield	Merino	0.58 ± 0.06	(Wuliji *et al.*, 2001)
	Merino	0.35 ± 0.05	(Mortimer & Atkins, 1989)
	Romney	0.40 ± 0.04	(Wuliji *et al.*, 2011)
	Corriedale	0.75 ± 0.15	(Benavides & Maher, 2000)
MFD	Merino	0.59 ± 0.06	(Wuliji *et al.*, 2001)
	Merino	0.47 ± 0.02	(Gifford *et al.*, 1995)
	Merino	0.48 ± 0.07	(Mortimer & Atkins, 1989)
	Romney	0.57 ± 0.05	(Wuliji *et al.*, 2011)
	Corriedale	0.65 ± 0.15	(Benavides & Maher, 2000)
	NZ-DP	0.40 ± 0.10	(Pickering *et al.*, 2013)
FDSD	Merino	0.51 ± 0.10	(Gifford *et al.*, 1995)
	NZ-DP	0.27 ± 0.11	(Pickering *et al.*, 2013)
CVFD	Merino	0.60 ± 0.06	(Wuliji *et al.*, 2001)
	NZ-DP	0.23 ± 0.10	(Pickering *et al.*, 2013)
MSS	Merino	0.39 ± 0.02	(Gifford *et al.*, 1995)
	Merino	0.13 ± 0.09	(Wuliji *et al.*, 2001)
	Romney	0.34 ± 0.14	(Hawker *et al.*, 1988)
	Romney	0.24 ± 0.05	(Wuliji *et al.*, 2011)
MSL	Merino	0.71 ± 0.11	(Wuliji *et al.*, 2001)
	Merino	0.44 ± 0.07	(Mortimer & Atkins, 1989)
	Romney	0.34 ± 0.14	(Hawker *et al.*, 1988)
	Romney	0.41 ± 0.06	(Wuliji *et al.*, 2011)
Curvature	NZ-DP	0.31 ± 0.10	(Pickering *et al.*, 2013)
Brightness	Merino	0.38 ± 0.06	(Wuliji *et al.*, 2001)
	Romney	0.23 ± 0.12	(Hawker *et al.*, 1988)
	Romney	0.12 ± 0.03	(Wuliji *et al.*, 2011)
	Corriedale	0.22 ± 0.11	(Benavides & Maher, 2000)
Yellowness	Merino	0.42 ± 0.07	(Wuliji *et al.*, 2001)
	Romney	0.14 ± 0.04	(Wuliji *et al.*, 2011)
	Corriedale	0.27 ± 0.13	(Benavides & Maher, 2000)

* NZ-DP: New Zealand dual-purpose sheep which are dominated by Romney, Coopworth, Perendale, Texel and composite crosses of these breeds.

$^{\Phi}$ SE: standard error.

1.2.3 Factors Affecting Wool Traits

Wool traits can be affected by many variables including breed of sheep, nutritional status, age, health status and shearing regime.

Merino sheep are the main source of high quality fine wool in New Zealand and Australia, while the Romney and its crosses of that breed produce the strongest wool clip in New Zealand.

Both the quantity and composition of feed has an effect on wool production. For example, either sulphur amino acids or high quality protein placed experimentally in the intestines of sheep, can achieve a large response in wool growth, and directly affect GFW and CFW (Reis, 1979). Wool production initially increases with age, reaching a maximum at three and a half years old before slowly declining. This means the age of individual sheep can affect GFW (Brown *et al.*, 1966). Various diseases and biological constraints including parasitic, bacterial and viral infections can also affect wool. For example, fleece rot, a common bacterial infection on the skin, can cause undesirable wool colouring. Finally, shearing time and frequency affects wool quality. Wool shorn in winter is typically of higher staple strength than wool shorn in summer. This is likely to be related to the proximity of the minimum fibre diameter in relation to the staple end. More frequently shorn wool has shorter staple lengths and higher clean scoured yields (Sumner & Hawker, 1986). Overall, wool traits are influenced by sheep genetics, environment and management practice (Cottle, 1991).

1.3 Follicles and the Mechanics of Wool Growth

Wool fibre is produced by the wool follicle in the skin of sheep in a similar way to how hair grows in other animals. The manner of growth within the follicle determines both the amount and characteristics of the fibre produced. It is largely determined by genetic factors (Rogers, 2006).

1.3.1 Follicle Structure

A mature follicle consists of an outer tube of cells, the outer root sheath (ORS), the inner root sheath (IRS) and a follicle bulb (Figure 1-1). In Zone 1, the cells in the follicle bulb undergo cell division and commence movement up the follicle. Differentiation initially consists of changes in size and shape. In Zone 2, keratin genes and keratin-associated protein (KAP) genes expression commences. Intermediate filaments and their matrix then start forming. In Zone 3, disulphide bonding, absorption and dehydration make the whole cellular structure becomes fixed and stable. In Zone 4, IRS degradation occurs (Zahn *et al.*, 2005).

Figure 1-1. The structure of a mature wool follicle. Sourced from Popescu and Höcker (2007).

1.3.2 Wool Follicle Development and Wool Growth

Wool follicles begin to form in the skin of the lamb during the last three months of gestation. In the surface layer of skin, cells aggregate to form the epidermal plug. These cells divide repeatedly and create a thin down-growth of epidermal cells into the lower skin. The accessory structures, including the arrector muscles, sweat and sebaceous glands grow out from the plug. After the epidermal plug ceases its downward growth, the cells at the plug base continue to divide and push back up towards to the surface. As these cells move toward the surface they differentiate into specialised parts of the wool fibre, and finally the fibre becomes mature and pushes up through the follicle to the skin surface (Popescu & Höcker, 2007).

The first follicles formed are the primary (P) follicles followed by secondary (S^o) follicles and then secondary-derived (S^D) follicles that branch from the S^o follicles (Figure 1-2). The branching of S^D can be extensive and determines 80% of the final follicle population density. The population density of P follicles does not differ widely between sheep breeds, while the population of total follicles varies between breeds of sheep and is genetically controlled (Hardy & Lyne, 1956). Selective breeding aims to reduce the difference in size between P and S^D follicles and increase the abundance of S^D follicles to result in fleeces of finer and more uniform fibre diameter.

"Material removed due to copyright compliance"

Figure 1-2. Diagram of wool follicles. (A) Primary wool follicle; (B) Secondary wool follicles; and (C) Secondary and secondary-derived follicles. Redrawn based on Rogers *et al.* (2006).

Wool is a proteinaceous material. Protein synthesis involves the activation of amino acids with adenosine triphosphate (ATP) which are then linked together to form polypeptides (proteins). It is thought that the protein synthesized takes the form of a random coils and these coils gradually become ordered into helical structures forming the fibrillar component of keratin (Ryder & Stephenson, 1968).

Follicle activity requires a lot of energy normally supplied by the way of glucose absorbed from the blood stream. In sheep, rumen-derived acetate is also an important source of energy (Leng & Stephenson, 1965). Glycose can be stored in the wool follicle as glycogen which can be readily

broken down to release glucose when it is required, mitigating reductions in fibre growth due to dietary fluctuations (Philpott & Kealey, 1991; Ryder & Stephenson, 1968).

Apart from wool proteins, the wool follicles also produce wax and suint (Henderson, 1968). Wool wax, containing lanolin, is not soluble in water and therefore cannot be washed out of the fleece by rain. It preserves fibre from adverse climatic effects. Having too little wool wax can leave the fibre 'weathered' at its tip; and having too much accounts for the low yield of clean fibre in some wools. Suint (or sweat) is soluble in water and can be easily washed away by rain. Suint plays an important part in the growth of populations of bacteria that can cause damage to the fleece (Simpson & Crawshaw, 2002).

1.3.3 Morphology of the Wool Fibre

The wool fibre consists of three major cellular structures: the cuticle, the cortex and the central medulla. Finer wool fibres generally do not have a medulla.

The cuticle is the external single-celled sheath, and it protects the fibre. The cells are rectangular (approximately 20 µm wide and 30 µm long), flattened (0.5 - 0.8 µm thick) and overlap each other to form a characteristic 'scale pattern' (Hocker, 2002). The degree of overlapping can be used to identify fibres from different breeds and species. Long-wool sheep have relatively long thin scales with little overlap, resulting in a smooth surface with a lustrous shine (Cottle, 1991). Each cuticle cell has a complex laminated structure comprised of three layers: the epicuticle (outermost layer), the exocuticle (middle layer) and the endocuticle (innermost layer) (Figure 1-3). The layers contain a variety of proteins and polysaccharides. The cuticle is in direct contact with the environment and plays a prominent role in the surface chemistry of the wool fibre.

The cortex comprises 90% of the fibre mass and is responsible for the major physical properties of wool. Corticle cells are spindle-shaped (45 - 95 µm long, 2 - 6 µm wide) and are essentially aligned along the length of the fibre (Hocker, 2002).The cortex usually consists of three types of cells: paracortex, orthocortex and mesocortex. Paracortex and orthocortex are defined as the two extremes, while mesocortex is an intermediary type (Figure 1-4).

The orthocortical cells are characterised by discrete nearly circular macrofibrils, often separated by cytoplasmic remnants and intermacrofibrillar material. Each macrofibril has a whorl-like appearance in cross-section (Figure 1-4) (Caldwell *et al.*, 2005; Marshall *et al.*, 1991; Rogers, 1959). In these cells the intermediate filament (IF)/matrix ratio is high. The paracortical cells have a small number of large macrofibrils, which are located around the centrally located cytoplasmic remnant. Their macrofibrils are so closely packed that is difficult to determine their boundary. The IFs show no distinctly ordered

Figure 1-3. Schematic diagram of a non-medulated wool fibre showing the major structural features. Modified from Marshall *et al.* (1991).

packing (Figure 1-4) and the IF/matrix ratio is low. From their microfibrilar structure, mesocortical cells appear to be an intermediate type of cells with characteristics that range from orthocortex-like to paracortex-like (Caldwell *et al.*, 2005; Marshall *et al.*, 1991; Rogers, 1959). At the ultra-structural level they are often recognized by a clearly resolved hexagonal or pseudo-hexagonal packing arrangement of IF in the macrofibrils (Figure 1-4) (Caldwell *et al.*, 2005). The IF/matrix ratio is also intermediate between that of the orthocortex and paracortex (Marshall *et al.*, 1991).

"Material removed due to copyright compliance"

Figure 1-4. Transmission electron micrographs of the cross-section of a Romney wool fibre. (A) A micrograph at a lower magnification showing three different types of cortical cells: paracortex (P), orthocortex (O) and mesocortex (M). Micrographs at a higher magnification showing the different IF packing modes of paracortical, mesocortical and orthocortical cells. Sourced from Caldwell *et al.* (2005).

In wool there is a considerable variation in the number and distribution of cell types both across transverse sections and along the length of the fibre. Orthocortical cells are the dominant type, generally making up over 50% of the fibre cross-section and volume (Marshall *et al.*, 1991). In addition, for samples from sheep from different breeds or within a breed, the proportion of orthocortical cells increases as the fibre diameter increases, while the proportion of mesocortical and

paracortical cells decreases, often with a partial substitution of the paracortex by the mesocortex. Highly crimped fibres, such as fine Merino wool, exhibit a well-defined bilateral arrangement of orthocortical and paracortical cells, with approximately half of the fibre being composed of each type (Horio & Kondo, 1953), and with the orthocortex on the convex side of the crimp wave. In poorly crimped fibres the paracortical cells are partly replaced by mesocortical cells or bilobate arrangements of paracortex occur (Orwin, 1989). It has been generally thought that the bilateral segmentation of orthocortex and paracortex is the cause of curvature, but there is still some debate about this (Marshall *et al.*, 1991). Amino acid analysis showed that the orthocortical cells contained more tyrosine, glycine, leucine and phenylalanine and less cysteine than the paracortical cells. Variation of the tension from the di-sulphide bonds may lead to various degree of crimp in the wool fibre (Campbell *et al.*, 1972).

In some of the medium and coarser wool, there is a central core called the medulla which runs lengthwise through the fibre. The proportion of the medulla varies greatly in wool fibre, ranging from 0 to 80% by volume of the whole fibre (Marshall *et al.*, 1991). Medulla size is likely to influence the rate at which wet processing treatments and dyes penetrate the fibre (Mishra, 2000).

1.3.4 Composition of the Wool Fibre

Wool fibre is primarily composed of proteins called hard α keratins. These have a high sulphur content as a consequence of the commonly occurring amino acid cysteine. In wool, the α-keratins are assembled into keratin intermediate filaments (KIFs), before being embedded in an inter-filamentous matrix consisting of KAPs (Popescu & Höcker, 2007). There is another protein called trichohyalin which is located in the inner root sheath and the medulla of the wool fibre, but it is not essential to fibre structure (Popescu & Höcker, 2007).

The earliest attempt to identify and classify wool proteins originated in 1935 (Goddard & Michaelis, 1935), and it divided the major wool components into two classes of extractable proteins: S-carboxy methyl kerateine A (SCMK-A) and S-carboxy methyl kerateine B (SCMK-B). They were lower and higher in sulphur content than the average sulphur content of wool respectively (Goddard & Michaelis, 1935). The SCMK-As became the hair (wool) keratins we know of today, while the SCMK-Bs represented the KAPs (Powell & Rogers, 1986).

1.4 Hair (Wool) Keratins

1.4.1 Keratin Structure

Keratins are fibrous structural proteins that form the IFs in epithelial cells. They are encoded by a large and well conserved gene family with 54 members identified in the human genome (Hesse *et al.*, 2001; Hesse *et al.*, 2004). Keratin proteins can be divided into two major types: the type I (acidic) keratins and type II (basic-neutral) keratins.

Keratin proteins show strong sequence conservation, especially in the central portion of the molecule (the so-called rod domain) (Steinert & Roop, 1988). The rod domain possesses a heptad sequence of the form $(a\text{-}b\text{-}c\text{-}d\text{-}e\text{-}f\text{-}g)_n$, with positions a and d commonly occupied by apolar resides, whereas other positions are often polar or charged residues (Fuchs & Weber, 1994; Steinert & Roop, 1988). This heptad sequence favours the formation of a coiled-coil in which right-handed α-helical coil around one another in a left-handed manner to form a rod-like structure (Fuchs & Weber, 1994; Steinert & Roop, 1988). Within this rod there are three breaks in the heptad continuity, designated by linkers L1, L12 and L2. There is an additional phasing discontinuity, known as a 'stutter' which results from the deletion of three residues from an otherwise continuous heptad structure, located close to the centre of segment 2B (Fuchs & Weber, 1994; Steinert & Roop, 1988). The head and tail domains of keratins are largely globular in nature, and often show sequence conservation among the members of each type. The head and tail domains can be further subdivided into end (E), variable (V) and homologous (H) subdomains (Steinert & Roop, 1988). The structural features of type I and type II keratin proteins are illustrated in Figure 1-5. Type II keratins possess a H1 subdomain of 36 amino acids and a H2 subdomain of 20 amino acids in the head and tail domain respectively, but the type I keratins only possess a short and variable H1 subdomain in the head domain, but without a H2 subdomain in the tail (Steinert & Roop, 1988). The subdomain E1 and V1 subdomains are usually larger, more basic and more glycine- and serine-rich than E2 and V2 subdomains.

1.4.2 Comparison of Epithelial and Hair Keratins

Both type I and type II keratins can be categorised into epithelial keratins and hair keratins. Epithelial keratins often possess head and tail domains rich in glycine and serine, but lack cysteine and have a low overall cysteine content. Hair keratins by comparison possess a high cysteine and proline content in their head and tail domains (Moll *et al.*, 2008; Rogers *et al.*, 2006).

Type I epithelial and type I hair keratins also have differences in their gene structures (Moll *et al.*, 2008; Rogers *et al.*, 1998; Rogers *et al.*, 2000; Rogers *et al.*, 2006). Type I epithelial keratin genes

possess eight exons, whereas type I hair keratin genes contain seven exons. However, both type II epithelial and hair keratin genes consist of nine exons.

The differences in protein properties for both types of keratins and in gene structures for type I keratins, as well as in the major locations of expression, have allowed further division of keratins from both types into epithelial and hair keratins (Moll *et al.*, 2008; Rogers *et al.*, 2006). The epithelial keratins are not directly relevant to this study and are therefore not discussed further in this review.

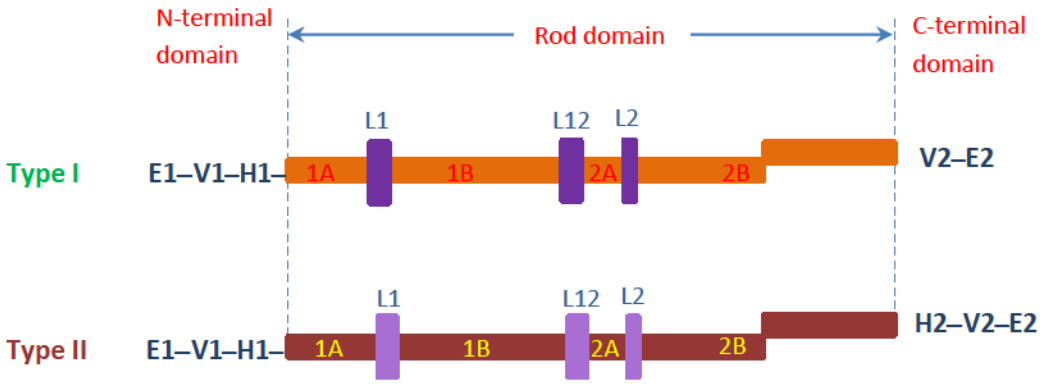


Figure 1-5. Schematic structure of the type I and type II hair keratins. Both proteins consist of non-helical end domains and a central α-helical rod domain. The rod domain is interrupted by three non-helical linkers L1, L12 and L2, and a break in the regularity of the heptad substructure near the middle of segment 2B. E is the end termini, V is a variable sequence region and H is a region of high sequence homology. Redrawn based on Steinert and Roop (1988).

1.4.3 Diversity within Hair (Wool) Keratins

Hair keratins have been most comprehensively studied in humans. Research in this species reveals 17 members of hair keratins, 11 type I and six type II (Langbein *et al.*, 2001; Langbein *et al.*, 2007; Langbein *et al.*, 1999). These are: K31, K32, K33A, K33B, K34 - K40 for type I, and K81- K86 for type II, under the new nomenclature proposed by Schweizer *et al.* (2006). The type II hair keratin genes are clustered on chromosome 12.q13.13, flanked by two epithelial keratin genes *KRT75* and *KRT6* (Rogers *et al.*, 2005). The type I hair keratin genes form two sub-clusters on chromosome 17q21.2, with *KRT39* and *KRT40* forming one sub-cluster and the others forming another sub-cluster (Langbein *et al.*, 2007; Rogers *et al.*, 2004a).

In wool, early electrophoretic analyses showed that wool keratins can be divided up into three main components called components 5, 7 and 8. Components 7 and 8 can be further separated into three and four sub-components respectively (Crewther *et al.*, 1975; Marshall, 1981). This suggests that there may be eight members of hair keratins in wool, four type I (components 8a, 8b, 8c1 and 8c2) and four type II (components 5, 7a, 7b and 7c). The proteins of four hair keratins have been isolated and their amino acid sequences have been determined, including two type I (component 8c1 or K31, and keratin A or K34) (Dowling *et al.*, 1986; Gough *et al.*, 1978) and three type II (component 5 or K85, component 7c or K83, and K86) hair keratins (Crewther *et al.*, 1978; Sparrow *et al.*, 1989; Sparrow *et al.*, 1992). More modern proteomic analyses on wool have identified ten hair keratins in wool, six type I and four type II (Plowman *et al.*, 2012). These keratins are K31, K33A, K33B, K34, K35 and K38 for type I, and K81, K83, K85 and K86 for type II, under the new nomenclature proposed by Schweizer *et al*. (2006).

DNA sequencing has revealed 17 hair keratin gene sequences in sheep, including ten for type I (*KRT31, KRT32, KRT33A, KRT33B, KRT34 - KRT36, KRT38 - KRT40*) and 7 for type II (*KRT81 - KRT87*) (Powell *et al.*, 1992; Powell *et al.*, 1993; Wilson *et al.*, 1988; Yu *et al.*, 2011). These hair keratins are shown in Table 1-3. This number is comparable to that identified in humans, suggesting that humans and sheep may have the similar number of hair keratins despite the relative hairlessness of humans. However, further investigations are required before such a conclusion can be drawn.

With regard to similarities between human and ovine keratin genes, it must be noted that since most of the hair keratin gene sequences are derived from cDNA sequences (Yu *et al.*, 2011), the gene sequences need to be physically mapped to the ovine chromosome to determine whether they represent different genes or different variants or isoforms of the same gene. Secondly, humans have a number of type I and type II hair keratin pseudogenes (Rogers *et al.*, 2004a; Rogers *et al.*, 2005), but further research is needed to determine whether the orthologs of these human pseudogenes are functional in sheep. It has been reported that human type I hair keratin pseudogene φhHaA has functional orthologs in the chimpanzee and gorilla (Winter *et al.*, 2001). Thirdly, bioinformatics analyses of the sheep genome sequence will allow the complete sheep hair keratin catalogue to be identified, especially with the continuing progress on the sheep genome sequence assembly.

While little is known about genetic variation in human hair keratin genes, there is some research in sheep revealing that the hair keratin genes are polymorphic (Itenge-Mweza *et al.*, 2007; Mckenzie *et al.*, 2012; McLaren *et al.*, 1997). To date, five, two and six alleles have been found for *KRT33A* (*KRT1.2*), *KRT83* (*KRT2.10*) and *KRT87* (*KRT 2.13*), respectively (Itenge-Mweza *et al.*, 2007; Mckenzie *et al.*, 2012; McLaren *et al.*, 1997).

Table 1-3. Hair keratins identified in sheep

Protein name	Gene name	Old name(s)	Sequence accession number[ϕ]	Reference
Type I keratin (10)				
K31	*KRT31*	Component 8c1; K1.1	P02534*	(Dowling *et al.*, 1986)
K32	*KRT32*		HQ283078, HQ283077	(Yu *et al.*, 2011)
K33A	*KRT33A*	47.6 kDa; K1.2	M23912	(Wilson *et al.*, 1988)
K33B	*KRT33B*		HQ283080	(Yu *et al.*, 2011)
K34	*KRT34*	Keratin A	0409251A*	(Gough *et al.*, 1978)
K35	*KRT35*		EU216426	(Yu *et al.*, 2011)
K36	*KRT36*		HQ283082	(Yu *et al.*, 2011)
K38	*KRT38*		EU216427	(Yu *et al.*, 2011)
K39	*KRT39*		HQ283083	(Yu *et al.*, 2011)
K40	*KRT40*		HQ283084	(Yu *et al.*, 2011)
Type II Keratin (7)				
K81	*KRT81*	K2.9	X62509	(Powell *et al.*, 1992)
K82	*KRT82*		HQ283086	(Yu *et al.*, 2011)
K83	*KRT83*	Component 7c; K2.10	P15241*	(Sparrow *et al.*, 1989)
K84	*KRT84*		HQ283088	(Yu *et al.*, 2011)
K85	*KRT85*	Component 5; K2.12	P25691*	(Sparrow *et al.*, 1992)
K86	*KRT86*	K2.11	P02539*	(Crewther *et al.*, 1978)
K87	*KRT87*	K2.13	X72379	(Powell *et al.*, 1993)

[ϕ] All sequence accession numbers refer to the GenBank, and the protein sequences from the GenBank Protein database are indicated with *.

1.4.4 Differential Keratin Expression within the Follicle

In humans, all hair keratin genes are expressed in the hair follicle apart from type II hair keratin gene *KRT84*. Keratin K84 is not detected in human hair, but is present in the filiform papillae of the tongue (Langbein *et al.*, 2001). However, *KRT84* is reported to be expressed in the sheep wool follicle, where a new type II hair keratin gene called *KRT87* is also expressed (Yu *et al.*, 2011).

Hair keratin genes exhibit differential and often sequential expression patterns in the hair follicle, resulting in distinct type I and type II hair keratin pairing in different hair cuticle and human matrix/cortex compartments (Langbein *et al.*, 2001; Langbein *et al.*, 2007; Langbein *et al.*, 1999). *KRT35* and *KRT85* are expressed to form type I K35 and type II K85 keratin pairs in the hair-forming matrix of the cortex and the cuticle. Due to the delayed onset of type II K82 expression in the lowermost hair cuticle, two type I keratins (K32 and K35) compete for type II keratin K85 in filament formation, but this situation is reversed in the mid-cuticle region, where two type II keratins (K82 and K85) are now found as opposed to only one type I keratin K32. In the cortex, the IFs originate essentially from the K31 and K85 keratin pair in the lower part, whereas higher up in the mid-cortex their individual composition cannot be predicted owing to the simultaneous expression of many type I and type II keratins. In contrast, the upper portions of the cortex and the cuticle each predominantly exhibit two type I keratins (K34/K39 and K39/K40, respectively) and one type II keratin (K86 and K82, respectively) during filament formation (Figure 1-6A).

The wool keratin genes are expressed in patterns that are similar to the human hair keratin genes overall, but some differences are also observed (Yu *et al.*, 2011) (Figure 1-6B). A notable difference is that there are more hair keratin genes from both type I and type I being expressed from the lower region through to the upper region of the wool follicle cortex. This results in an increased possibility of type I and type II keratin pairing in the cortex.

1.5 Keratin-associated Proteins (KAPs)

Keratin-associated proteins are the proteins that form the matrix between the keratin IFs. KAPs are thought to surround the keratin IFs during development in the follicle and interact with IFs through extensive disulphide bond cross-linking between cysteine residues in the head and tail domains of hair keratins (Powell & Rogers, 1997). Bundles of keratin IFs combine with KAPs to form macrofibrils through inter- and/or intra-molecular disulphide bond formation (Rogers, 2004). While KAPs may have little or no discernable effect on keratin IF structure, their effect on keratin IF assembly into larger arrays is considered to be crucial (Plowman, 2003). It is therefore believed that KAPs play a critical role in defining the physico-mechanical properties of the hair and wool fibres.

KAPs are relatively small in size (ca. 10 - 30 kDa) and characteristically possess either a high cysteine or high glycine and tyrosine content (Powell and Rogers, 1997; Rogers *et al.*, 2006). Originally they were categorised into three main groups according to their amino acid composition: the high sulphur (HS; ≤30 mol% cysteine), the ultra-high sulphur (UHS; >30 mol% cysteine) and the high glycine/tyrosine (HGT; 35 - 60 mol% glycine and tyrosine) groups (Powell & Rogers, 1997).

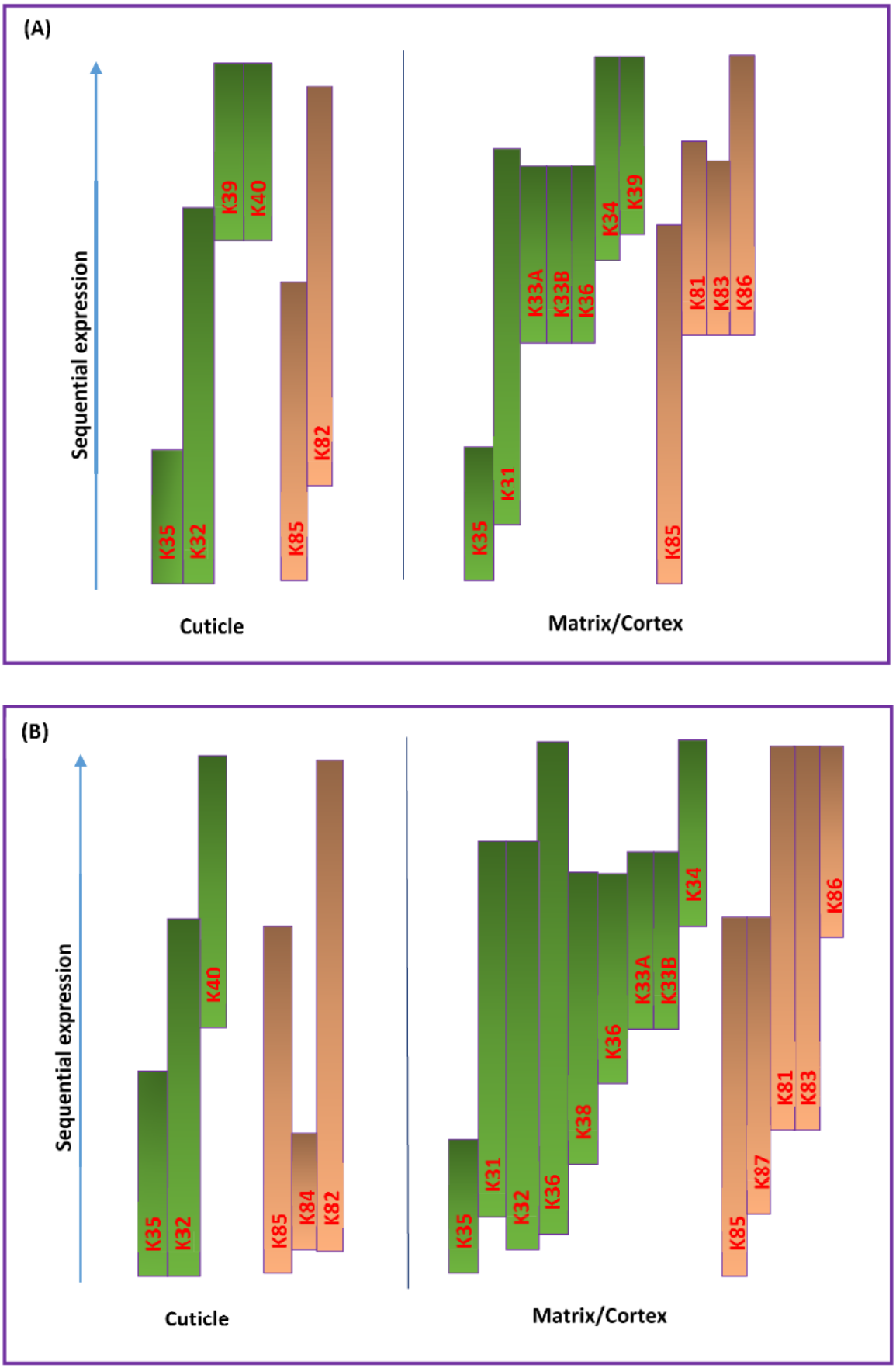


Figure 1-6. Schematic presentation of the expression patterns of hair keratin genes in the human hair (A) and wool (B) follicles. Redrawn from Langbein *et al.* (2007) and Yu *et al.* (2011).

1.5.1 The Diversity of KAPs

Two-dimensional gel electrophoretic analyses on wool SCMK-Bs showed that a large number of HS, UHS and HGT KAPs existed (Powell & Rogers, 1997). Eight KAP proteins (B2A or KAP1-1, B2B or KAP1-2, B2C or KAP1-3, BIIIA3A or KAP2-1, BIIIA3 or KAP2-3, BIIIB2 or KAP3-1, BIIIB3 or KAP3-3 and BIIIB4 or KAP3-3) from the HS group were isolated and their amino acid sequences determined (Elleman, 1971, 1972a, 1972b; Elleman & Dopheide, 1972; Haylett & Swart, 1969; Haylett *et al.*, 1971; Swart & Haylett, 1971; Swart & Haylett, 1973). Two HGT proteins were also isolated and characterised by amino acid sequencing (Dopheide, 1973; Gillespie, 1990). With the advent of DNA cloning techniques, 16 gene sequences encoding the HS (Frenkel *et al.*, 1989; Powell *et al.*, 1983), UHS (Fratini *et al.*, 1994; Jenkins & Powell, 1994; MacKinnon *et al.*, 1990) and HGT (Fratini *et al.*, 1993; Kuczek & Rogers, 1985; Kuczek & Rogers, 1987) KAPs have been discovered in sheep through the isolation of their cDNA or gene sequences. Prior to this study, a total of 27 protein and DNA sequences had been identified (Table 1-4). Sequence analyses indicate that these sequences probably represent 21 different KAP members, but the possibility cannot be ruled out that some KAP members may represent variants of the same gene. This has been reported for human KAP1 genes. Some KAP1 gene sequences that were previously thought to represent different genes were subsequently found to be allelic variants of a single gene (Shimomura *et al.*, 2002). A similar problem was also reported for the human KAP4 gene (Kariya *et al.*, 2005).

Recently, our understanding of the complexity of KAP has advanced significantly with the advent of the large-scale whole-genome sequencing of human KAP genes and the completion of the human genome sequence. Extensive bioinformatics analyses have revealed a total of 88 functional and 17 pseudo-KAP genes in the human genome (Rogers & Schweizer, 2005; Rogers *et al.*, 2008; Rogers *et al.*, 1998; Rogers *et al.*, 2002; Rogers *et al.*, 2004b; Rogers *et al.*, 2001; Rogers *et al.*, 2007; Shibuya *et al.*, 2004). This probably represents the complete number of KAPs in the human genome.

1.5.2 The Categorisation of KAPs

KAPs are assigned into families based on amino acid sequence similarity. In general, individual KAP families can be distinguished by the content of cysteine or glycine-tyrosine, the type and number of repeat structures, and occurrence of unique sequence motifs. Unique motifs are often found in individual families (Rogers *et al.*, 2006).

The 88 functional KAP genes found in humans have been assigned into 25 KAP families: KAP1 to KAP27, with the exception of KAP14 and KAP18 which are only found in mouse. These families include 12 HS (KAP1-KAP3, KAP11, KAP13, KAP15-KAP16 and KAP23-KAP27), four UHS (KAP4, KAP5,

Table 1-4. Ovine KAP protein and gene sequences identified.

KAP family	KAP member	Old name	Sequence type	Sequence accession number[Φ]	Reference
KAP1	KAP1-1	B2A	Protein*	P02438	(Elleman, 1972b)
		B2A	DNA	X01610	(Powell *et al.*, 1983)
	KAP1-2	B2B	Protein	P02439	(Elleman & Dopheide, 1972)
	KAP1-3	B2C	Protein	711148A	(Elleman, 1971)
		B2C	DNA	X02925	(Powell *et al.*, 1983)
	KAP1-4	B2D	DNA	X01610	(Powell *et al.*, 1983)
KAP2	KAP2-1	BIIIA3A, KAP2.12	Protein	P02443	(Swart & Haylett, 1973)
	KAP2-3	BIIIA3	Protein	P02441	(Swart & Haylett, 1973)
			DNA*	U60024	Unpublished
KAP3	KAP3-1A	BIIIB2	Protein	P02446	(Haylett & Swart, 1969)
	KAP3-1B	BIIIB2	DNA	M21099	(Frenkel *et al.*, 1989)
	KAP3-2	BIIIB3	Protein	P02444	(Haylett *et al.*, 1971)
			DNA	M21100	(Frenkel *et al.*, 1989)
	KAP3-3	BIIIB4	Protein	P02445	(Swart & Haylett, 1971)
			DNA	M21103	(Frenkel *et al.*, 1989)
KAP4	KAP4-1		DNA*	X73462	(Fratini *et al.*, 1994)
	KAP4-2		Unknown		(Powell & Rogers, 1997)
	KAP4-3		DNA	EU239778	(Yu *et al.*, 2009)
KAP5	KAP5-1		DNA	X55294	(MacKinnon *et al.*, 1990)
	KAP5-2		DNA*		(Powell & Rogers, 1986)
	KAP5-4		DNA	X73434	(Jenkins & Powell, 1994)
	KAP5-5		DNA *	X73435	(Jenkins & Powell, 1994)
KAP6	KAP6-1	HGT type II	DNA	M95719	(Fratini *et al.*, 1993)
	KAP6-2	HGT type II	Protein		(Gillespie, 1990)
KAP7	KAP7-1	HGT-C2	DNA	X05638	(Kuczek & Rogers, 1987)
KAP8	KAP8-1	HGT-F	Protein	P02448	(Dopheide, 1973)
		HGT-F	DNA	X05639	(Kuczek & Rogers, 1987)

* Partial sequence

[Φ] All sequence accession numbers refer to the GenBank, except for 711148A which refers to the EMBL.

KAP9, KAP10, KAP12 and KAP17), and seven HGT (KAP6-KAP8 and KAP19-KAP22) KAP families. Assignment of these human KAP genes into families has been described and reviewed elsewhere (Rogers *et al.*, 2006; Rogers *et al.*, 2002; Rogers *et al.*, 2004b; Rogers *et al.*, 2001; Rogers *et al.*, 2007; Shibuya *et al.*, 2004; Yahagi *et al.*, 2004), so will not be covered in this review.

The 19 individual KAP members reported in sheep can be assigned into eight families: KAP1 to KAP8 (Table 1-4). KAP1 to KAP3 are HS-KAP families, KAP4 and KAP5 are UHS families, whereas KAP6 to KAP8 are HGT-KAP families. The assignment of these wool KAP members into families is described briefly below:

KAP1: There are probably four members of KAP1 present in sheep (KAP1-1, KAP1-2, KAP1-3 and KAP1-4), although the gene encoding KAP1-2 was not described until research was undertaken for this thesis. This number matches that found in humans (Rogers *et al.*, 2006; Rogers *et al.*, 2001). KAP1 proteins are highly conserved and differ mainly in the number of tandem decapeptide repeats QTSCCQP(T/I)(S/C)(I/L) in the N-terminal half of the proteins (Figure 1-7a).

KAP2: There are two protein sequences (BIIIA3A and BIIIA3) reported in this family. These two sequences share 95% homology. BIIIA3A is referred to as the ortholog of human KAP2-1 (Rogers et al., 2006), although the same sequence was previously called KAP2.12 (Powell and Rogers, 1997). BIIIA3 is identical to the predicated amino acid sequence of a partial DNA sequence referred to as KAP2-3 (GenBank U60024). To date no ovine KAP2-2 gene has not been reported. Given that the KAP2 family is highly conserved in humans with over 97% sequence identity between the family members (Rogers *et al.*, 2006; Rogers *et al.*, 2001), it is likely that these two wool KAP2 sequences represent two members of the same family. KAP2 family members possess several cysteine-rich pentameric repeat structures (CCXPX) and length differences are also observed among family members (Figure 1-7b).

KAP3: There are three major (designated as BIIIB2, BIIIB3 and BIIIB4) and one minor (designated as BIIIB1) proteins reported for wool KAP3 (Swart *et al.*, 1969). Only three proteins have been isolated and sequenced (Haylett & Swart, 1969; Haylett *et al.*, 1971; Swart & Haylett, 1971). The three gene sequences thought to encode the major proteins have also been identified (Frenkel *et al.*, 1989). These members are named KAP3-1, KAP3-2 and KAP3-3 respectively for BIIIB2, BIIIB3 and BIIIB4. Minor sequence differences between the protein sequence and the gene sequence for each member are observed, which may reflect nucleotide variation in the gene or sequencing errors. However, the BIIIB2 protein sequence (labelled as KAP3-1A) appears to be different to the predicted amino acid sequence of the BIIIB2 gene (labelled as KAP3-1B), sharing only 94% homology in amino acid sequence. It is unknown whether these two sequences may represent two different members.

However, the human KAP3 family has three functional genes and one pseudogene (Rogers *et al.*, 2006; Rogers *et al.*, 2001). KAP3-2 and KAP3-3 are very similar to each other, but different to KAP3-1 (Figure 1-7c). Little repetitiveness of structure is seen in this family.

KAP4: KAP4 is a UHS-KAP family. There are three members identified in sheep (Fratini *et al.*, 1994; Powell & Rogers, 1997; Yu *et al.*, 2009). KAP4 proteins possess a rigid repeat structure that covers a large portion of the mid-section of the protein, consisting of concatenates of monocysteine- and dicysteine-containing pentameric repeats (Figure 1-7d).

KAP5: KAP5 belongs to the UHS group and is one of the largest KAP families, with possibly 12 family members present in humans (Rogers & Schweizer, 2005; Rogers *et al.*, 2006). In sheep, three completed and one partial DNA sequences of this KAP family have been identified (Jenkins & Powell, 1994; MacKinnon *et al.*, 1990; Powell & Rogers, 1986), and designated as KAP5-1, KAP5-2, KAP5-4 and KAP5-5. Sheep KAP5 proteins possess repeat structures of either cysteine/serine-rich or glycine-rich, and the repeat numbers vary among members (Figure 1-7e).

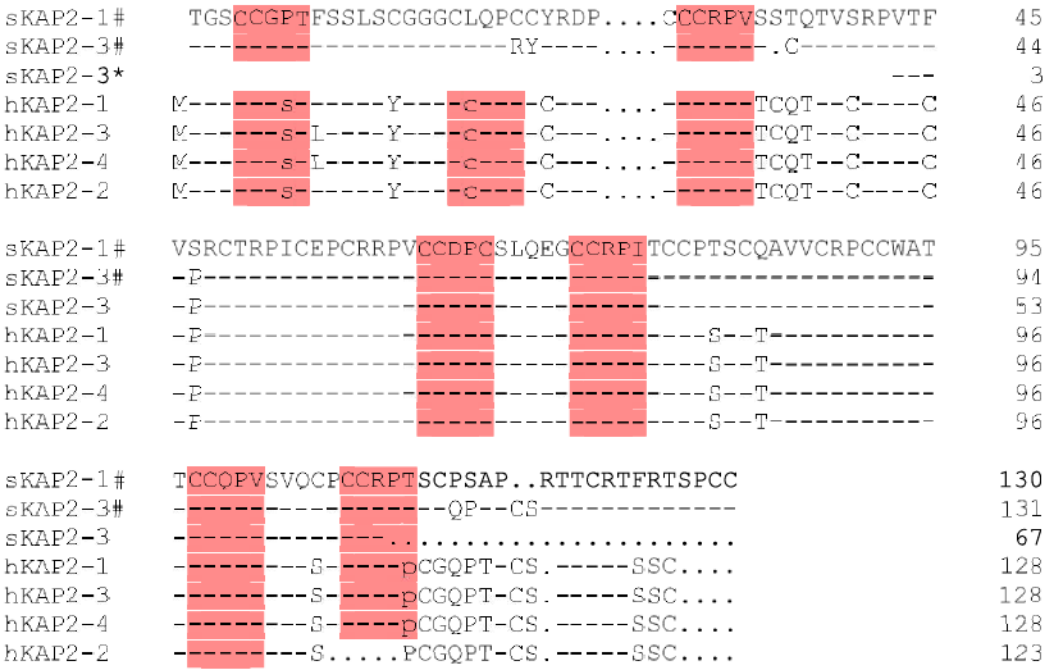
KAP6: KAP6 is a HGT-KAP family. Three gene members have been identified in the human genome (Rogers *et al.*, 2002). Southern-hybridisation analysis revealed that there are several members in sheep and as many as twenty in mice (Fratini *et al.*, 1993). In sheep, only one DNA sequence (Fratini *et al.*, 1993) and one partial protein sequence (Gillespie, 1990) are available. KAP6 proteins consist of repetitive units of glycine-tyrosine and glycine-tyrosine-glycine (Figure 1-7f).

KAP7 and KAP8: Both KAP7 and KAP8 belong to the HGT group. These two KAP families are thought to contain only one family member each. There is high sequence similarity between orthologs from sheep and humans, and the number of glycine-tyrosine repeat structures is small compared to the KAP6 family (Figure 1-7g and h).

(a)


sKAP1-1 MACCSTSFCGFPICSTGGTCGSSPCQPTCCQTSCCQPTSIQTSCCQPISI 50
sKAP1-2# SV CG S 49
sKAP1-3# A C-RS S 39
sKAP1-3 A C-RS S 40
sKAP1-4 T NF T 50
sKAP1-1 QTSCCQPTSI QTSCCQPTCLQTSGCETGCGIGGSIGYGQV 90
sKAP1-2# D 79
sKAP1-3# 69
sKAP1-3 T 70
sKAP1-4 QTSCCQPISI 100
sKAP1-1 GSSGAVSSRTRWCRPDCRVEGTSLPPCCVVSCTPPSCCQLYYAQASCCRP 140
sKAP1-2# S 129
sKAP1-3# S 119
sKAP1-3 S 120
sKAP1-4 K S 150
sKAP1-1 SYCGQSCCRPACCCQPTCIEPICEPSCCEPTC 172
sKAP1-2# V 156
sKAP1-3# T V T-SQ-I 151
sKAP1-3 T V T-SQ-I 152
sKAP1-4 V T 182

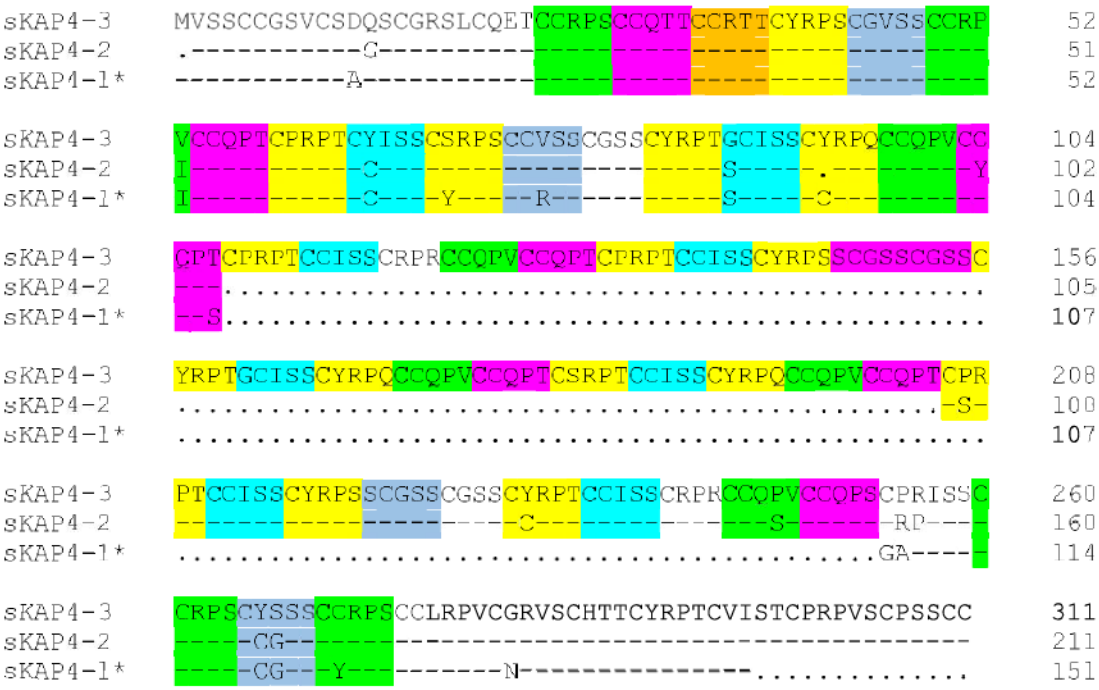
(b)


sKAP2-1# TGSCCGPTFSSLSCGGGCLQPCCYRDP CCCRPVSSTQTVSRPVTF 45
sKAP2-3# RY .C 44
sKAP2-3* 3
hKAP2-1 s Y c C TCQT C 46
hKAP2-3 s L Y c C TCQT C 46
hKAP2-4 s L Y c C TCQT C 46
hKAP2-2 s Y c C TCQT C 46
sKAP2-1# VSRCTRPICEPCRRPVCCDPCSLQEGCCRPITCCPTSCQAVVCRPCCWAT 95
sKAP2-3# P 94
sKAP2-3 P 53
hKAP2-1 P S T 96
hKAP2-3 P S T 96
hKAP2-4 P S T 96
hKAP2-2 P S T 96
sKAP2-1# TCCQPVSVQCPCCRPTSCPSAP..RTTCRTFRTSPCC 130
sKAP2-3# QP CS 131
sKAP2-3 67
hKAP2-1 S pCGQPT-CS. SSC 128
hKAP2-3 S pCGQPT-CS. SSC 128
hKAP2-4 S pCGQPT-CS. SSC 128
hKAP2-2 S PCGQPT-CS. SSC 123

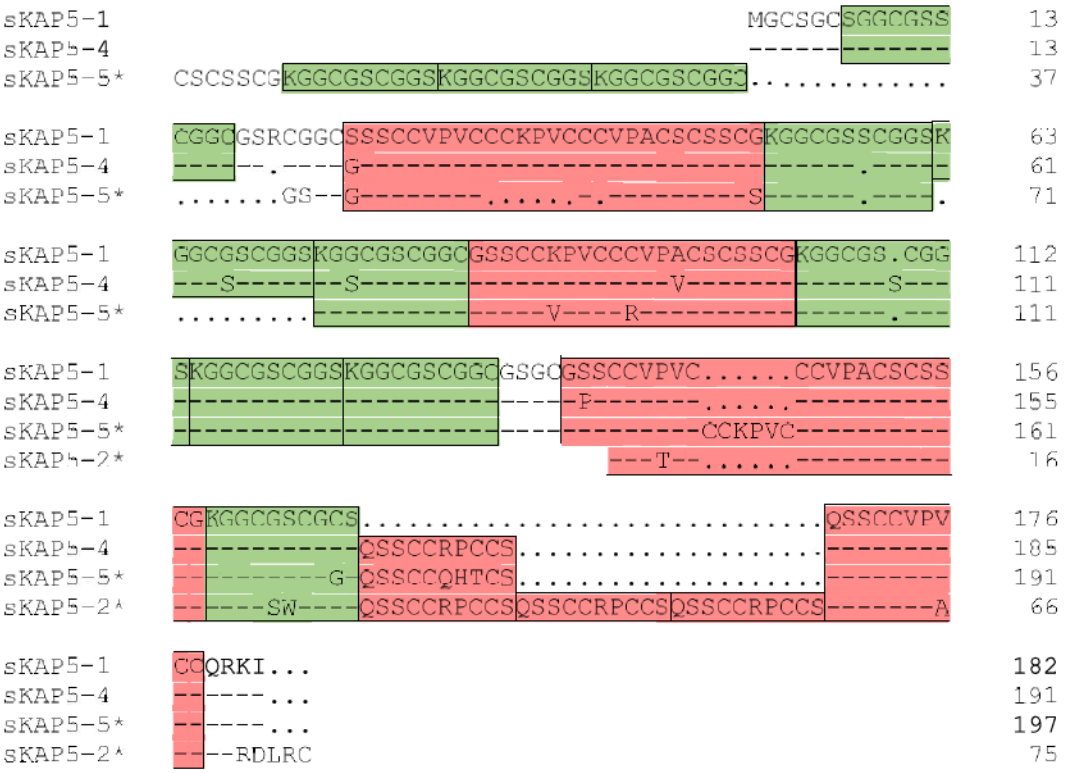
(c)


sKAP3-1A# ACCAPRCCSVRTGPATTICSSDKFCRCGVCLPSTCPHNISLLQPTCC.D 48
sKAP3-1B M----------S---D---------.- 49
hKAP3-1 MY---L-S---P------F--F--S----------E------I--.- 49
sKAP3-2* ----------tvw--e----.- 25
sKAP3-2# ----RL----P-S----------TVW-------C- 49
sKAP3-3# ----RL----P-S----------TVWF------C- 49
sKAP3-3 M----RL----P-T----------TVWF-E----.- 49
hKAP3-2 MD---S-S---P----------S----------TVW--E-I--.- 49
hKAP3-3 MD---S-G---P----------S----------TVW--E----.- 49
sKAP3-1A# NSPVPCVYPDTYVPTCFLLNSSHPTPGLSGINLTTFIQPGCENVCEPRC 97
sKAP3-1B ------YV----------R----------R--------- 98
hKAP3-1 TC-P--CK--------W---NC----------YV-----SP----- 98
sKAP3-2 -R-P-YHV-QPS----------Q-----ES--H--YT-SS--PCIPSC- 74
sKAP3-2# -R-P-YHV-QPS----------Q-----ES-----YT-SS--PCIPSC- 98
sKAP3-3# -R-P--HI-QPS----------Q-----ES-----YT--S--PCIPSC- 98
sKAP3-3 -R-P--HI-QPS----------Q-----ES-----YT--S--PCIPSC- 98
hKAP3-2 -C-P--HI-QPC---------CQ-----ETL-----T--C--PCLPRG- 98
hKAP3-3 -C-P--HI-QPC---------CQ-----ETL-----T--C--PCLPRG- 98

(d)


sKAP4-3 MVSSCCGSVCSDQSCGRSLCQETCCRPSCCQTTCCRTTCYRPSCGVSSCCRP 52
sKAP4-2 .----------C---------- 51
sKAP4-1* ----------A---------- 52
sKAP4-3 VCCQPTCPRPTCYISSCSRPSCCVSSCGSSCYRPTGCISSCYRPQCCQPVCC 104
sKAP4-2 I----------C----------S-----.---------Y 102
sKAP4-1* I----------C----Y------R----------S-----C---------- 104
sKAP4-3 CPTCPRPTCCISSCRPRCCQPVCCQPTCPRPTCCISSCYRPSSCGSSCGSSC 156
sKAP4-2 ---..... 105
sKAP4-1* --S..... 107
sKAP4-3 YRPTGCISSCYRPQCCQPVCCQPTCSRPTCCISSCYRPQCCQPVCCQPTCPR 208
sKAP4-2 .....-S- 100
sKAP4-1* 107
sKAP4-3 PTCCISSCYRPSSCGSSCGSSCYRPTCCISSCRPRCCQPVCCQPSCPRISSC 260
sKAP4-2 ----------C----------S----------RP---- 160
sKAP4-1* .....GA----- 114
sKAP4-3 CRPSCYSSSCCRPSCCLRPVCGRVSCHTTCYRPTCVISTCPRPVSCPSSCC 311
sKAP4-2 -----CG---------- 211
sKAP4-1* -----CG---Y---------N----------..... 151

(e)



(f)

(g)

```
sKAP7-1     MTRFFCCGSYFPGYPSYGTNFHRTFRATPLNCVVPLGSPLGY..GCNGYS      48
hKAP7-1     ---Y----------I------G----------------N-GC------      50

sKAP7-1     SLGYGFGGSSFSNLGCGYGGSFYRPWGSGSGFGYSTY                  85
hKAP7-1     ----S----NIN---GC--------------------                  87
```

(h)

```
sKAP8-1     M.SYCFSSTVFPGCYWGSYGYPLGYSVGCGYGSTYSPVGYGFGYGYNGSG     49
hKAP8-1     -LCDN-PGA-------------------------------------C-      50

sKAP8-1     ASGCRRFWPFALY                                          62
hKAP8-1     -F-Y--YS-----                                          63
```

Figure 1-7. Amino acid alignments of the ovine KAP sequences from individual KAP families together with representatives of human sequences. (a) KAP1 family; (b) KAP2 family; (c) KAP3 family; (d) KAP4 family; (e) KAP5 family; (f) KAP6 family; (g) KAP7 family; and (h) KAP8 family. Dashes represent amino acids identical to the top sequence, and dots have been introduced to improve alignments. Different repeat sequences are highlighted in different colours, except for the GY and GYG repeats seen in the HGT-KAPs. The protein sequences are indicated by #, and the others are the predicated amino acid sequences from DNA sequences. * indicates partial sequences. The ovine KAPs are indicated with a prefix of "s" whereas the humans are with a prefix of "h". Two human KAP2 members (hKAP2-1A and hKAP2-1B) have an identical amino acid sequence, and are therefore shown as hKAP2-1. Accession numbers or references for ovine sequences are shown in Table 1-4, and those for human sequences are: AC007455 (hKAP2-1A); AC037482 (hKAP2-1B); AC025904 (hKAP2-2, hKAP2-3 and hKAP2-4); AC007455 (hKAP3-2 and hKAP3-2); AP001708 (hKAP6-1, hKAP6-2 and hKAP6-3); AB096962 (hKAP7-1); and AP001709 (hKAP8-1).

1.5.3 The Clustering of KAP genes

The detection of multiple KAP genes from one single genomic clone from a variety of mammalian species suggests the grouping of KAP genes in the genome (Cole & Reeves, 1998; Kuhn *et al.*, 1999; MacKinnon *et al.*, 1990; Powell *et al.*, 1983). Recent bioinformatics analyses of the human genome sequence have confirmed the clustering of KAP genes (Rogers *et al.*, 2002; Rogers *et al.*, 2004b; Rogers *et al.*, 2001; Shibuya *et al.*, 2004; Yahagi *et al.*, 2004).

In humans, all the KAP genes are clustered into five chromosome domains: 17q21.2 (containing KAP1 to KAP4, KAP9, KAP16 and KAP17), 21q22.1 (KAP6 to KAP8, KAP11, KAP13, KAP15, KAP19 to KAP26), 21q22.3 (containing KAP10 and KAP12), 11p15.5 (containing some KAP5 genes) and 11q13.4 (containing other KAP5 genes) (Rogers *et al.*, 2002; Rogers *et al.*, 2004b; Rogers *et al.*, 2001; Shibuya *et al.*, 2004; Yahagi *et al.*, 2004).

In sheep, the precise location of the KAP genes identified is currently unknown, but seven KAP genes are reported to be mapped to three chromosomes. These are chromosome 1 (containing *KRTAP6-1, KRTAP7-1* and *KRTAP8*-1), chromosome 11 (containing *KRTAP1-1, KRTAP1-3* and *KRTAP3-2*), and chromosome 21 (containing *KRTAP5-1*) (McLaren *et al.*, 1997; Parsons *et al.*, 1994a; Wood *et al.*, 1992).

1.5.4 Polymorphism within KAP Genes

While the KAP genes have been well characterised in humans, limited effort has been made to understand the full scenario of variation within genes. To date, variation has only been investigated in the KAP1 and KAP4 families and these studies were restricted to Caucasian and Japanese populations (Kariya *et al.*, 2005; Shimomura *et al.*, 2002). Four previously identified KAP1 genes (*KRTAP1-1A, KRTAP1-1B, KRTAP1-6* and *KRTAP1-7*) have been confirmed to be allelic variants, and another four previously reported KAP1 genes (*KRTAP1-8A, KRTAP1-8B, KRTAP1-3* and *KRTAP1-9*) are also allelic variants of another KAP1 member (Shimomura *et al.*, 2002). In the KAP4 family, ten of the eleven members exhibit polymorphism, each having two or three allelic variants (Kariya *et al.*, 2005).

In sheep, variation has been investigated in *KRTAP1-1, KRTAP1-3, KRTAP3-2, KRTAP6-1, KRTAP7-1* and *KRTAP8-1*, and all the genes investigated exhibit varying degrees of polymorphism (Itenge-Mweza *et al.*, 2007; McLaren *et al.*, 1997; Rogers *et al.*, 1994b). Up to nine alleles have been reported for *KRTAP1-3* (Itenge-Mweza *et al.*, 2007). The more extensive polymorphism found in sheep compared with humans is possibly due to a greater number of sheep samples being screened.

Variation detected in *KRTAPs* includes both single nucleotide polymorphisms (SNPs) and length polymorphisms. Variation in length is notable in both sheep and human KAP1 genes (Rogers *et al.*, 1994b; Shimomura *et al.*, 2002), and human KAP4 genes (Kariya *et al.*, 2005). Length polymorphisms appear to be the result of having a variable number of cysteine-rich repeated coding sequences. These have probably arisen by intragenic deletion and/or duplication events of the repeated segments during evolution (Kariya *et al.*, 2005).

1.5.5 The Expression of KAPs

In humans, all the KAP genes identified are expressed in the hair follicle, except for families 16, 22, 25 and 27 (Rogers *et al.*, 2008; Rogers *et al.*, 2006; Rogers *et al.*, 2007). These KAP genes exhibit a uniform hair follicle expression pattern that is characterised by the chromosome domain in which these gene are located. The exception is for the KAP genes on chromosome domain 21q22.1 and a single KAP gene (*KRTAP17-1*) on chromosome domain 17q21.2 (Rogers *et al.*, 2006) (Figure 1-8).

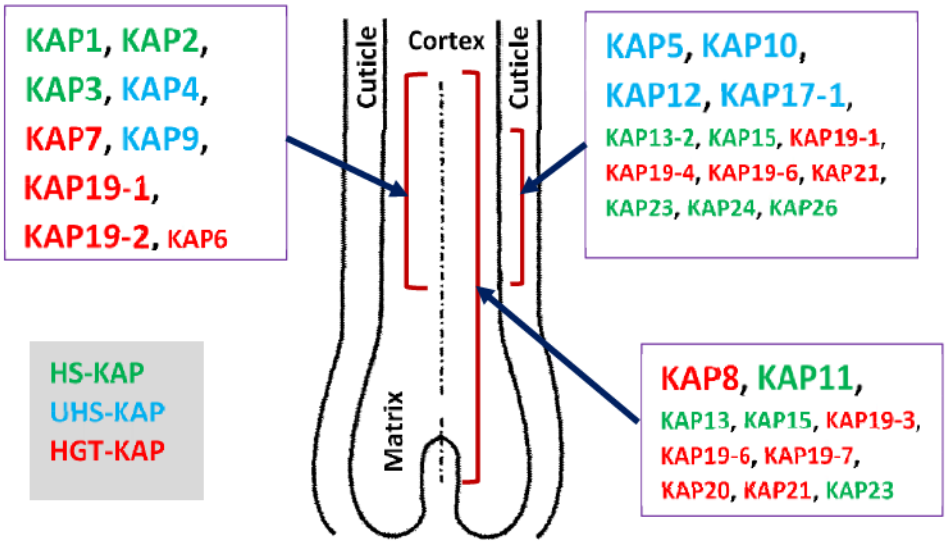


Figure 1-8. Scheme of KAP gene expression in the human hair follicle. The level of gene expression is indicated by the size of the KAP lettering, with the larger letter size representing strong expression and the smaller letter size representing weak expression. Redrawn from Rogers *et al.* (2006).

The expression of KAP genes on chromosome domain 17q21.2 except *KRTAP17-1* (i.e. *KRTAP1-n, KRTAP2-n, KRTAP3-n, KRTAP4-n* and *KRTAP9-n*) occurs high up in the differentiated portion of the hair cortex (Rogers *et al.*, 2006; Rogers *et al.*, 2001), a region where a loss of water and an absence of free sulphydryl groups associated with hair fibre cross-linking was previously shown (Powell &

Rogers, 1997). All the KAP genes on chromosome domains 21q22.3, 11p15.5 and 11q13.4 are expressed high up in the hair cuticle (Rogers *et al.*, 2006; Rogers *et al.*, 2004b; Yahagi *et al.*, 2004).

However, the expression of KAP genes (*KRTAP6-n, KRTAP7-1, KRTAP8-1, KRTAP11-1, KRTAP13-n, KRTAP15-1, KRTAP19-n, KRTAP20-n, KRTAP23-1, KRTAP24-1* and *KRTAP26-1*) on chromosome domain 21q22.1 shows varying degrees and locations of expression in the hair follicle (Rogers *et al.*, 2008; Rogers *et al.*, 2006; Rogers *et al.*, 2002). While most of the KAP genes on this chromosome domain are seen in the cortex and matrix cell region, some genes are also expressed in the hair cuticle such as *KRTAP13-2, KRTAP19-n* and *KRTAP23-1* (Rogers *et al.*, 2006; Rogers *et al.*, 2004b). Other genes are expressed exclusively in the differentiated region of the hair cuticle, such as *KRTAP19-4*, *KRTAP24-1* and *KRTAP26-1* (Rogers *et al.*, 2008; Rogers *et al.*, 2006; Rogers *et al.*, 2007). High levels of expression is observed for the HGT genes *KRTAP7-1, KRTAP8-1, KRTAP19-1*, *KRTAP19-2* and the HS gene *KRTAP11-1*, whereas the other KAP genes such as *KRTAP6-n* are weakly expressed (Rogers *et al.*, 2006; Rogers *et al.*, 2004b). A schematic summary of human KAP expression is shown in Figure 1-8.

The expression data of KAPs is relatively limited in sheep. Of the eight KAP families identified in sheep, the HGT-KAPs (KAP6, KAP7 and KAP8) are the first expressed KAPs and appear in the orthocortical cells. The HS-KAPs (KAP1, KAP2 and KAP3) are initially expressed in the orthocortical cells, but soon after are expressed in all cortical cells (Powell & Rogers, 1997). The UHS-KAP4 family is expressed slightly later and only in paracortical cells (Fratini *et al.*, 1994). The UH-KAP5 family is expressed late in hair cuticle differentiation (Jenkins & Powell, 1994; MacKinnon *et al.*, 1990). A schematic summary of wool KAP expression is shown in Figure 1-9.

The expression of the KAP genes seen in sheep is consistent overall with that reported in humans, but some differences are also observed. The most notable is the further division of expression sites seen in the wool follicles. For example, KAP1 to KAP4, KAP6 to KAP7 are all expressed in the hair cortex, but in the wool follicle, KAP6 and KAP7 are only expressed in the orthocortex and the expression of KAP4 is restricted to the paracortex. While KAP1 and KAP2 are expressed in the whole cortex of wool follicles, they start from the orthorcortical half. Such a well-defined bilateral structure is not found in human hair follicle. In humans, KAP8 is expressed earlier than KAP6 and KAP7, but they are co-expressed in the wool follicle.

The weak expression of KAP6 in human hair is of interest as there is no evidence that KAP6 is weakly expressed in the wool follicle. It should be noted that the level of expression reported for the human KAP genes is based on the quantitation of mRNA level and not on the protein level. There is only a moderate correlation between mRNA transcript levels and protein translation levels (Miklos & Maleszka, 2001), and the level of mRNA expression is insufficient to predict protein expression levels,

especially for genes with low mRNA expression. It has been reported that for some genes with similar levels of mRNA expression, the protein levels exhibit over a 20-fold and as much as a 30-fold difference (Gygi *et al.*, 1999).

Figure 1-9. Scheme of keratin and KAP gene expression sites in the wool follicle. Redrawn from Powell & Rogers (1997).

1.6 The Potential for the Development of Gene Markers for Wool Traits

1.6.1 Traditional Methods for Genetic Selection for Wool Quality

Traditionally, selective sheep breeding has assisted in improving both the volume and quality of wool fibre produced and has been shown to accumulate desirable "wool genes" in the population over time. With the use of quantitative genetic techniques, genetically superior parents are identified and assigned an estimated breeding value (eBV) which describes the future potential of that animal as a dam or sire for a particular wool trait. However, wool traits, like many other production traits, do not exhibit simple inheritance patterns and will exhibit quantitative variation in phenotype depending on environment. The eBVs are therefore of limited benefit in assessing the true genetic merit of sheep. An added complication of using eBVs is that wool characteristics are only fully expressed when a sheep is mature. Typically one needs to wait at least three years to reliably assess the wool

characteristics of an individual sheep before it can be reliably used in a traditional selective breeding programme. Genetic flock improvement can therefore be slow and inefficient.

In contrast to traditional breeding approaches, the use of genetic markers can be more accurate and efficient at selecting superior animals. Genetic markers are specific DNA sequences that contribute to, or influence a trait. The genetic marker breeding approach allows earlier assessment of an animal's potential and more accurate control over the genetic variation in the family pedigree.

1.6.2 Genetic Marker Technology

Genetic markers are specific nucleotide sequence differences that contribute to, or influence, a trait. A significant advantage of genetic markers is that they can identify genetic merit for traits that are difficult to measure on a live animal or cannot be measured early in life, and that they are not reliant on phenotypic expression for measurement (Williams, 2005). They are accordingly useful in improving the accuracy of genetic selection.

There are several ways to identify genetic markers, but the most commonly used approaches are the genome-wide scanning or linkage mapping approach (Bush & Moore, 2012; Hirschhorn & Daly, 2005), and the candidate gene approach (Kwon & Goate, 2000; Zhu & Zhao, 2007), each having advantages and disadvantages.

In the genome-wide scanning approach, the whole genome, which may contain a large number of candidate genes, is searched to identify quantitative trait loci (QTL). Without any presuppositions regarding the importance of specific functional features of the investigated traits, it can be both expensive and resource intensive (Bush & Moore, 2012; Hirschhorn & Daly, 2005). In the candidate gene approach, known genes that are thought to be responsible for the phenotypic variance of a trait are directly targeted for investigation (Kwon & Goate, 2000; Zhu & Zhao, 2007). It is a more effective and economical approach, although one "only detects what one is looking for".

The candidate gene approach is better suited for detecting genes underlying common and more complex traits. However, one of the limitations of this approach is its 'hit and miss' nature (Zhu & Zhao, 2007). Its success relies mainly on two factors. The first is choosing suitable candidate genes that play a relevant role in the process or disease under investigation, and the second is that it requires an accurate and effective method to detect variation occurring in the candidate genes (Kwon & Goate, 2000; Zhu & Zhao, 2007).

1.6.3 Techniques for Detecting Genetic Variation

Numerous techniques for detecting DNA-level variation have been developed and they can be grouped into two categories based on the purpose of detection:

- Diagnostic techniques which are only able to detect known polymorphism, such as allele-specific PCR (David & Deutch, 1992), allele-specific oligonucleotide hybridisation (Conner *et al.*, 1983), and the fluorogenic 5' nuclease (TaqMan) PCR assay (Livak, 1999);

- Screening techniques which are capable of detecting both known and new polymorphism, such as DNA sequencing (Engelke *et al.*, 1988), PCR-restriction fragment length polymorphism (PCR-RFLP) (Tatari *et al.*, 1995), PCR-denaturing gradient gel (PCR-DGGE) analysis (Myers *et al.*, 1985), heteroduplex analysis (HDA) (White *et al.*, 1992), PCR-single-strand conformational polymorphism (PCR-SSCP) analysis (Orita *et al.*, 1989) and PCR-stem-loop conformational polymorphism (PCR-SLCP) (Zhou *et al.*, 2011). Among these, PCR-SSCP has become a popular technique for detecting polymorphism and genotyping samples due to its simplicity and sensitivity (Hayashi, 1991, 1992; Miterski *et al.*, 2000).

1.6.4 Previous Findings on KAP Genes Associated with Variation in Wool Traits

There have been a number of studies describing associations between wool traits and variation in KAP genes.

Wool crimp is thought to be affected by the composition of KAPs in the wool fibre. The felting lustre (FL) mutant wool contains little HGT-KAPs (Gillespie & Darskus, 1971). Examination of transcript prevalence revealed that the KAP6-1, KAP7-1 and KAP8-1 genes are down-regulated, while the KAP2-12 and KAP4-2 genes are up-regulated in mutant follicles (Li *et al.*, 2009). In the mutant follicle, there is only one type of cortical cells (the paracortical cells) and the orthocortical cells where the HGT-KAPs are expressed are absent (Li *et al.*, 2009).

There have been some previous studies describing an association between variation in *KRTAPs* and wool traits. Parson *et al.* (1994b) reported that variation in a KAP6 gene was associated with MFD in medium wool Peppin Merinos. The KAP6 gene maps to ovine chromosome 1 (Parsons *et al.*, 1994a) where Beh *et al*. (2001) detected a QTL affecting MFD in medium wool Merinos. However such an association was not confirmed by Roldan *et al.* (2010).

Rogers *et al*. (1994a) reported a putative QTL for wool staple strength on ovine chromosome 11 in the region spanning *KRTAP1-1*, *KRTAP1-3* and *KRT33A* (*KRT1-2*) in Romney sheep. On this

chromosome, Roldan et al. (2010) identified a QTL for wool weights, staple strength and coefficient of variation of fibre diameter.

Recently variation in *KRTAP1-1* was reported to be associated with various wool traits (Itenge-Mweza, 2007). Itenge-Mweza (2007) described variation in *KRTAP1-1* associated with wool yield in one half-sib family and to be associated with MSL and wool brightness in another half-sib family.

These findings suggest that KAP genes potentially affect wool traits and further investigation is required to explore this potential.

1.7 Aim and Objectives of this Study

KAPs are a major structural component of hair/wool fibre in the matrix and cross-link with the keratin intermediate filaments, the other main structural component. Given that there are many KAP genes and all the KAP genes are likely polymorphic and potentially expressed, it is hypothesized that variation in KAP genes affects the expression and/or the structure of KAPs thereby affecting wool traits.

The primary aim of this study was to characterise genetic variation in ovine KAP genes and to investigate whether this variation affects wool traits. This will potentially enable a better understanding of the genetic basis for variation in wool traits and potentially, possibly leading to the development of gene-markers for breeding sheep for better quality wool.

While the KAP genes have been almost completely characterised in humans, only a small number of genes has been identified in sheep and the majority of human KAP orthologs remain to be characterised in this species. In addition, variation has only been investigated in seven of the 18 functional KAP genes identified in sheep. A better understanding the KAP genes and their variation in this species will provide foundation for association studies. The first objective of this study is therefore to identify new KAP genes and investigate variation in these newly identified and previously identified KAP genes.

To accommodate the extensive diversity seen in KAP genes, the current nomenclature also required revision. The second objective of this study was to propose an updated nomenclature for KAP genes.

The final objective was to investigate whether variation in KAP genes was associated with specific wool traits, and thus to ascertain whether gene-markers can be developed to improve wool quality.

Chapter 2

Identification of Ovine Keratin-associated Protein (KAP) Genes and Investigation of Variation in those Genes

2.1 Introduction

The KAPs are a complex class of proteins and are encoded by a large number of genes that appear to be polymorphic. More than 100 genes arranged in at least 27 families have been identified in mammals (Rogers & Schweizer, 2005; Rogers *et al.*, 2008; Rogers *et al.*, 2007), but in sheep, only 13 functional KAP genes have been identified to date. These are *KRTAP1-1* (Powell *et al.*, 1983), *KRTAP1-3* (Powell *et al.*, 1983), *KRTAP1-4* (Powell *et al.*, 1983), *KRTAP3-1* (Frenkel *et al.*, 1989), *KRTAP3-2* (Frenkel *et al.*, 1989), *KRTAP4-1* (Fratini *et al.*, 1994), *KRTAP4-3* (Fratini *et al.*, 1994), *KRTAP5-1* (Jenkins & Powell, 1994), *KRTAP5-4* (Jenkins & Powell, 1994), *KRTAP5-5* (Jenkins & Powell, 1994), *KRTAP6-1* (Fratini *et al.*, 1993), *KRTAP7-1* (Kuczek & Rogers, 1987) and *KRTAP8-1* (Kuczek & Rogers, 1987). Homologues for many of the genes described in humans remain to be identified in the sheep genome, but the recent release of a draft ovine genome sequence will enable this to be investigated.

While the ovine KAP genes appear to be polymorphic, genetic variation has only been reported in seven of the genes including *KRTAP1-1* (Rogers *et al.*, 1994b; McLaren *et al.*, 1997; Itenge-Mweza *et al.*, 2007), *KRTAP1*-3 (Rogers *et al.*, 1994b; McLaren *et al.*, 1997; Itenge-Mweza *et al.*, 2007), *KRTAP3-2* (McLaren *et al.*, 1997), *KRTAP5-1* (McLaren *et al.*, 1997), *KRTAP6-1* (McLaren *et al.*, 1997), *KRTAP7-1* (McLaren *et al.*, 1997) and *KRTAP8-1* (Wood *et al.*, 1992). The limited knowledge of ovine KAP genes and variation in those genes has hampered research into whether these KAP genes affect wool traits.

In this context, the first part of this chapter provides details on an investigation into variation in some previously identified KAP genes: specifically *KRTAP1-4, KRTAP5-4, KRTAP6-n, KRTAP7-1* and *KRTAP8-1*. To date there are no reports of variation in *KRTAP1-4, KRTAP5-4* and *KRTAP6-n*, but variation in *KRTAP7-1* and *KRTAP8-1* has been described previously (Wood *et al.*, 1992; McLaren *et al.*, 1997). Variation in *KRTAP7-1* was identified using Southern hybridisation-RFLP, but no sequence information was revealed (McLaren *et al.*, 1997), while variation in *KRTAP8-1* was only described in the promoter region (Wood *et al.*, 1992) and not in the coding region.

The second part of this chapter describes the identification of new ovine KAP genes and variation that is present in them. These include *KRTAP1-2, KRTAP8-2, KRTAP11-1, KRTAP13-3* and *KRTAP24-1*.

While a protein sequence (Swiss-Prot P02439.1) for ovine KAP1-2 (also known as B2B or SCMK-B2B) had been known for nearly four decades (Elleman & Dopheide, 1972), the gene (notionally called *KRTAP1-2*) encoding ovine KAP1-2, had not been identified. This raised the question as to whether the ovine KAP1-2 protein sequence actually represented allelic variation, or an RNA processing artefact, of one of the three known KAP1 genes (*KRTAP1-1, KRTAP1-3* or *KRTAP1-4*).

KAP8 was thought to be a single-gene family in sheep (Kuczek & Rogers, 1987) and humans (Rogers *et al.*, 2002). Recently, a new KAP8 gene called *KRTAP8-2* was identified in goats (Jin *et al.*, 2011). This suggests that the number of KAP8 genes may vary between species and, given the relatedness of sheep and goats, a second member of the KAP8 family may exist in sheep.

Bovine sequences for the putative *KRTAP11-1* and *KRTAP13-3* had been reported with GenBank accession numbers NM_001080740 and ENSBTAG00000040032, respectively. These two genes were searched for in sheep, together with a homologue for *KRTAP24-1*. Human *KRTAP24-1* encodes a KAP of low cysteine content when compared to the majority of known human KAP family members (Rogers *et al.*, 2007).

2.2 Materials and Methods

2.2.1 Sheep Blood Samples and DNA Purification

Gene identification and screening to identify variation was carried out in 250-320 not closely related sheep from a variety of breeds, including Merino, New Zealand Romney, Coopworth, and cross-bred sheep. Blood samples from these sheep were collected onto FTA cards (Whatman BioScience, Middlesex, UK), and genomic DNA used for PCR amplification was purified using a two-step washing procedure as described in Zhou *et al.* (2006).

2.2.2 Search of KAP Homologues in the Sheep Genome Sequence

The ovine KAP1-2 gene was identified using an 87 bp nucleotide sequence to BLAST against the Ovine Genome Sequence Assembly v1.0 (www.livestockgenomics.csiro.au/sheep) (v2.0 was not available when the study was undertaken). This 87 bp sequence was identified to be conserved across the ovine KAP1-1, KAP1-3 and KAP1-4 genes, and it notionally would translate to produce a sequence found in the ovine KAP1-2 protein sequence (Swiss-Prot P02439.1). This was presumed to be the notional KAP1-2 gene.

The ovine KAP24-1 gene was identified using the coding sequence of the human KAP24-1 gene (NM_001085455) (Rogers *et al.*, 2007) to BLAST search the Ovine Genome Assembly v2.0 (Livestock

Genomics, 2011). The genome sequence that showed the highest homology with the human *KRTAP24-1* sequence was presumed to be the notional ovine KAP24-1 gene.

The ovine KAP8-2 gene was searched for using the coding sequence of the caprine KAP8-2 gene (GenBank AY510123) to BLAST against the Ovine Genome Assembly v2.0 (Livestock Genomics, 2011). The ovine sequence that showed the highest homology with the caprine *KRTAP8-2* sequence was presumed to be the notional ovine KAP8-2 gene.

The ovine assembly sequences identified above were used to design PCR primers for amplifying the individual genes.

2.2.3 PCR Primers and Amplification

PCR primers for amplifying *KRTAP1-2*, *KRTAP24-1* and *KRTAP8-2* (Table 2-1) were designed based on the sheep genome assembly sequences described above. Based on the *KRTAP1-2* sequence obtained in this study, two further internal primers were subsequently designed for amplifying a shorter PCR amplicon suitable for SSCP analysis. The *KRTAP24-1* primers contained an 11-mer adapter (5'-gagactgactc-3') at the 5' end, which was added to enable PCR-SLCP analysis (Zhou *et al.*, 2011).

Primers for amplifying *KRTAP11-1* and *KRTAP13-3* (Table 2-1) were designed based on the bovine *KRTAP11-1* (NM_001080740) and *KRTAP13-3* (ENSBTAG00000040032) sequences, respectively.

The published ovine *KRTAP1-4, KRTAP5-4, KRTAP6-1, KRTAP7-1* and *KRTAP8-1* sequences were used to design primers (Table 2-1) for amplifying these individual genes.

The sequences of all these primers are shown in Table 2-1. The primers were synthesized by Integrated DNA Technologies (Coralville, IA, USA).

PCR amplifications for each gene were performed in a 20 µL reaction containing the genomic DNA on one 1.2 mm punch of FTA paper, 0.25 µM of each primer, 150 µM of each dNTP (Eppendorf, Hamburg, Germany), 2.5 mM of Mg^{2+}, 0.5 U of *Taq* DNA polymerase (Qiagen) and 1× reaction buffer supplied. The thermal profile consisted of 2 min at 94 °C, followed by 35 cycles of 30 s at 94 °C, 30 s at the annealing temperature shown in Table 2-1 and 30 s at 72 °C, with a final extension of 5 min at 72 °C. The extension time for *KRTAP1-2* (909 bp) was 60 s at 72 °C. Amplification was carried out in either iCyclers (Bio-Rad, Hercules, CA, USA) or S1000 thermal cyclers (Bio-Rad, Hercules, CA, USA).

Amplicons were visualised by electrophoresis in 1% agarose gels (Quantum Scientific, Queensland, Australia), using 1× TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM Na_2EDTA) containing 200 ng/mL of ethidium bromide.

Table 2-1. PCR primer sequences, annealing temperatures and SSCP/SLCP conditions used in this chapter for amplifying ovine KAP genes

Gene	Primer name	Primer sequence (5-3')*	Annealing temp	Extension time	Amplicon size[ϕ]	SSCP/SLCP conditions[†]	Reference sequence
KRTAP1-2	KRTAP1-2-up	TGAATGCTAATGACTACTTGAC	62 °C	60 s	909 bp		Ovine Assembly v1.0
	KRTAP1-2-dn	TATTGAATGGAATCAGATAAGTC					
	KRTAP1-2s-up	TAACAACCCTCCTCTCAATCT	62 °C	50 s	557 bp	7%, 220 V, 25 °C	This study
	KRTAP1-2s-dn	TTCATGGACTGAAGTTGAACT					
KRTAP1-4	KRTAP1-4-up	ATCCTCCAAGCATTACAATTC	57 °C	50 s	638 bp	7%, 220 V, 25 °C	X01610
	KRTAP1-4-dn	ATTCAGAGATACTGTGTCTTG					
KRTAP5-4	KRTAP-5-up	CTGCTCCTCTGACCTACTC	61 °C	50 s	678 bp	8%, 250 V, 25 °C	X73434
	KRTAP5-4-dn	CAGACAGCCTCACAGATGC					
KRTAP6	KRTAP6-up	TCTACCCGAGAACAACCTC	59 °C	50 s	317 bp	14%, 250 V, 25 °C	M95719
	KRTAP6-dn	AGGTATAGAGGATGAGAGTC					
KRTAP7-1	KRTAP7-up	ACTTGCTCTTCACATTCTATC	60 °C	30 s	327 bp	14%, 250 V, 25 °C	X05638
	KRTAP7-dn	GTAGTCATCTGGAGCCATG					
KRTAP8-1	KRTAP8-up	CATTCCCTGCTCTCCAAGC	60 °C	30 s	258 bp	14%, 250 V, 25 °C	X05639
	KRTAP8-dn	GAGAAGATTCCATGCCTCTG					
KRTAP8-2	KRTAP8-2-up	TAGGCAGTCAGTCATCCTG	60 °C	30 s	473 bp	14%, 200 V, 25 °C	Ovine Assembly v2.0
	KRTAP8-2-dn	ATAGAGAATATGAAGTCCACG					
KRTAP11-1	KRTAP11-1-up	TGCATCTCTCAACCAGCAC	61 °C	50 s	532 bp	7%, 250 V, 10 °C	NM_001080740
	KRTAP11-1-dn	TGGAATCTTGATTCACTCATG					
KRTAP13-3	KRTAP13-3-up	TACATTCAAACTCAGAATCTTC	60 °C	40 s	594 bp	8%, 280 V, 15 °C	ENSBTAG00000040032
	KRTAP13-3-dn	TGAATTTGGCTCTTCTACAAG					
KRTAP24-1	KRTAP24-up	gagactgactcAAATGTTGCCATGTTATTGTCC	60 °C	80 s	1003 bp	10%, 320 V, 14 °C	Ovine Assembly v2.0
	KRTAP24-dn	gagactgactcGACAAATGAAGTAATTCAGCAG					

* The 11-mer adapter sequence used in SLCP is shown in lower case.

[ϕ] Expected size based on the reference sequence.

[†] SLCP conditions are shaded.

2.2.4 Variation Screening

Variation in PCR amplicons was screened by SSCP analysis for all genes except *KRTAP24-1*. For *KRTAP24-1*, a SLCP analysis was used to detect variation in PCR amplicons, as SLCP gave better resolution than SSCP for this amplicon.

Both SSCP and SLCP were carried out using the same procedure. A 0.7 μL aliquot of each PCR amplicon was mixed with 7 μl of loading dye (98% formamide, 10 mM EDTA, 0.025% bromophenol blue, 0.025% xylene-cyanol). After denaturation at 95 °C for 5 min, samples were rapidly cooled on wet ice and then loaded on 16 cm × 18 cm acrylamide:bisacrylamide (37.5:1) (Bio-Rad) gels. Gel concentrations and electrophoresis conditions for the individual genes are shown in Table 2-1. All electrophoresis was performed using Protean II xi cells (Bio-Rad) for 18 h in 0.5 × TBE buffer. Gels were silver-stained according to the method of Byun *et al.* (2009).

2.2.5 Sequencing of PCR Amplicons

The 909 bp PCR amplicons of *KRTAP1-2* were directly sequenced without SSCP analysis, as sequence variation in such a large fragment cannot be detected by SSCP.

For all other PCR amplicons, amplicons representative of different SSCP or SLCP patterns were selected for DNA sequencing. Amplicons that appeared to be homozygous were subject to sequencing directly, however, for heterogeneous amplicons or the alleles that were only found in heterozygous sheep, individual sequences or alleles were firstly separated by SSCP or SLCP, and single bands of interest were recovered from gels as gel slices, macerated, and used as a template for re-amplification with the original primers. This second amplicon was then directly sequenced. Each PCR amplicon was sequenced in both directions until an identical sequence was obtained from at least two independent PCR reactions. All sequencing was carried out at the Lincoln University DNA Sequencing Facility.

2.2.6 Sequence Analysis

Sequence alignments, translations and comparisons were carried out using DNAMAN (version 5.2.10, Lynnon BioSoft, Vaudreuil, Canada). The BLAST algorithm was used to search the GenBank databases for homologous sequences. Prediction of the notional start codon was performed using the NetStart server (www.cbs.dtu.dk/services/NetStart). Potential phosphorylation sites were predicted using the NetPhos 2.0 Server (www.cbs.dtu.dk/services/NetPhos/). N-linked glycosylation sites were predicted using the NetNGlyc 1.0 Server (www.cbs.dtu.dk/services/NetNGlyc/), and O-linked glycosylation sites were predicted using the NetOGlyc 3.1 Server (www.cbs.dtu.dk/services/NetOGlyc/).

2.3 Results

2.3.1 Variation in Previous Described Ovine KAP Genes

Nine allelic variants detected for ovine *KRTAP1-4*

This work has been published in "Gong H, Zhou H and Hickford JGH. Polymorphism of the ovine keratin-associated protein 1-4 (KRTAP1-4) gene. Molecular Biology Reports 2000, 37:3377-80".

Amplicons of the expected size (638 bp) were obtained from all sheep blood samples using the *KRTAP1-4* primers designed in this study. These amplicons exhibited polymorphism upon SSCP analysis and nine unique banding patterns could be identified (Figure 2-1). Either one or a combination of two patterns was observed for each sheep.

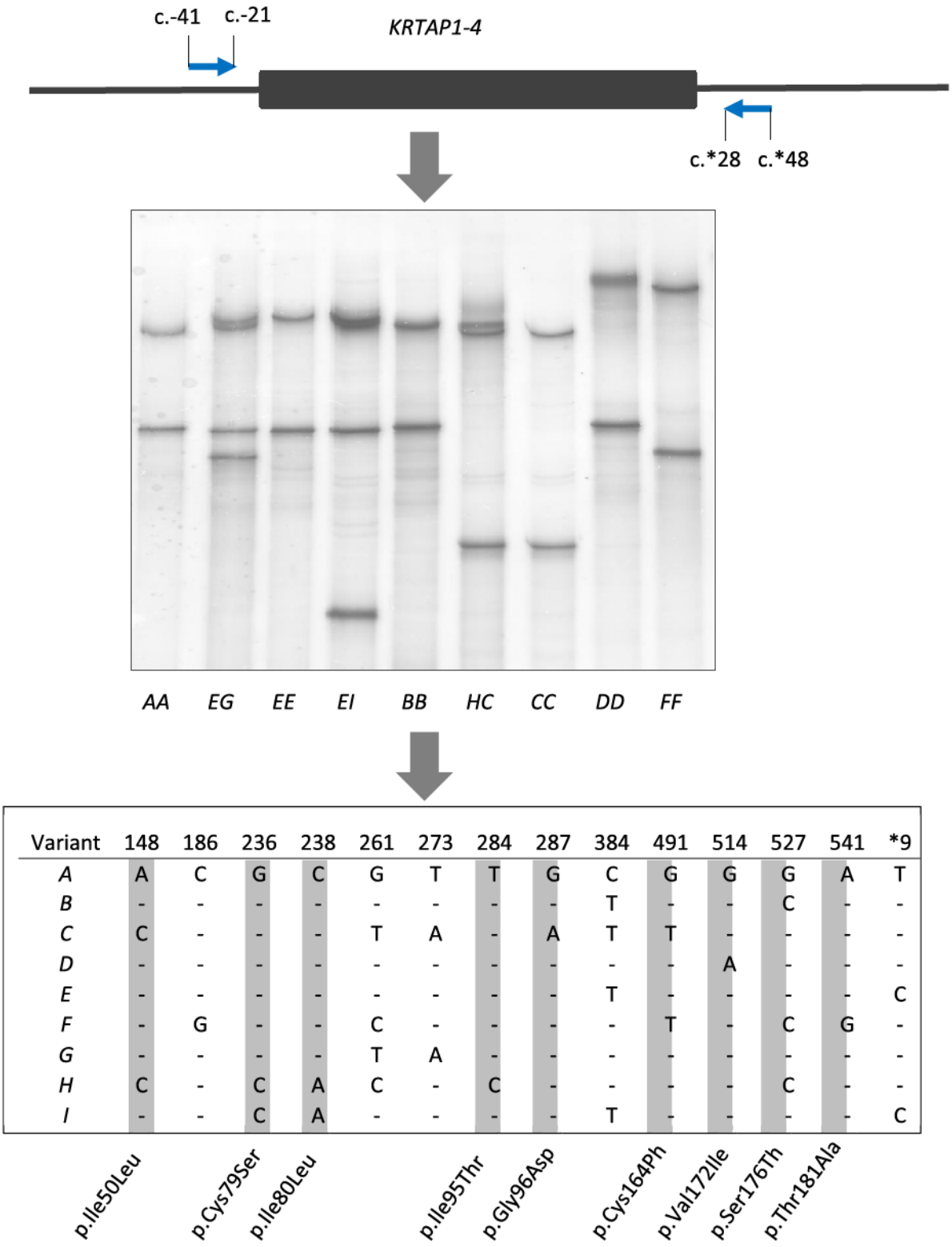
Sequencing of PCR amplicons representative of these unique SSCP patterns revealed nine unique nucleotide sequences. All of these sequences showed high similarity to the published ovine *KRTAP1-4* sequence (GenBank X01610). These sequences were named alleles *A* to *I* and deposited into the GenBank with the accession numbers GQ507741 - GQ507749. In the sheep typed, these alleles were detected at frequencies of 32.2%, 16.3%, 14.4%, 17.5%, 10.4%, 3.2%, 1.9%, 3.0% and 1.1% for *A* to *I*, respectively.

There were a total of 14 SNPs identified among the nine sequences, with 13 of them located in the coding region. Of these 13 SNPs, nine were non-synonymous SNPs that would result in amino acid changes (Figure 2-1).

Five allelic variants in ovine *KRTAP5-4*

This work has been published in "Gong H, Zhou H, Plowman J, Dyer J and Hickford JGH. Analysis of variation in the ovine ultra-high sulphur keratin-associated protein KAP5-4 gene using PCR-SSCP technique. Electrophoresis 2010, 31:3545-7".

Using the PCR primers designed for the ovine KAP5-4 gene, all of sheep blood samples generated PCR amplicons of the expected size (approximately 678 bp). These amplicons exhibited polymorphism upon SSCP analysis and five unique banding patterns could be identified (Figure 2-2). Either one or a combination of two patterns was observed for each sheep.



Variant	148	186	236	238	261	273	284	287	384	491	514	527	541	*9
A	A	C	G	C	G	T	T	G	C	G	G	G	A	T
B	-	-	-	-	-	-	-	-	T	-	-	C	-	-
C	C	-	-	-	T	A	-	A	T	T	-	-	-	-
D	-	-	-	-	-	-	-	-	-	-	A	-	-	-
E	-	-	-	-	-	-	-	-	T	-	-	-	-	C
F	-	G	-	-	C	-	-	-	-	T	-	C	G	-
G	-	-	-	-	T	A	-	-	-	-	-	-	-	-
H	C	-	C	A	C	-	C	-	-	-	-	C	-	-
I	-	-	C	A	-	-	-	-	T	-	-	-	-	C

p.Ile50Leu p.Cys79Ser p.Ile80Leu p.Ile95Thr p.Gly96Asp p.Cys164Ph p.Val172Ile p.Ser176Th p.Thr181Ala

Figure 2-1. Variation identified in ovine *KRTAP1-4*. The entire coding sequence of *KRTAP1-4* was amplified with PCR using primers KRTAP1-4-up and KRTAP1-4-dn with the location of primers indicated. Nine unique PCR-SSCP banding patterns representing nine allelic variants (*A* to *I*) in either homozygous or heterozygous genotypes are shown. Fourteen SNPs were identified and the non-synonymous SNPs are shaded with the corresponding amino acid substations being shown. Nucleotide positions refer to GenBank X01610 and the numbering of nucleotide and amino acids follow the HGVS nomenclature (den Dunnen & Antonarakis, 2000).

Sequencing of PCR amplicons representative of the unique SSCP patterns revealed five different nucleotide sequences (Figure 2-2). One sequence (named *A*) was identical to the published ovine *KRTAP5-4* sequence (GenBank X73434), whereas the remaining four sequences (named *B* to *E*) were unique, but showed high sequence similarity to the published ovine *KRTAP5-4* sequence (GenBank X73434). The newly identified allele *B* (c.517_546dup), allele *C* (c.494G>C), allele *D* (c.[70G>A; 412G>T; 546C>T; 592G>A]) and allele *E* (c.[546C>T; 566G>A] are available in GenBank with the accession numbers GU255997 - GU256001. The sequence variations that are shown (Figure 2-2) record the difference between that the *A* allele (X73434) and the other alleles. These alleles were detected at frequencies of 51.8%, 11.2%, 8.4%, 15.8% and 12.8% for *A, B, C, D* and *E*, respectively, in the sheep typed.

Sequence analysis revealed six SNPs and one length polymorphism in the ovine KAP5-4 gene (Figure 2-2). Of the six SNPs identified, five were located in the coding region and four of them were non-synonymous SNPs that would potentially result in amino acid changes. The length polymorphism occurred in the 3' coding sequence, with sequences having either one or two repeats of a 30 nucleotide sequence encoding a cysteine-rich decapeptide "RPCCSQSSCC".

Five *KRTAP6* sequences from three putative loci

This work has been published in "Gong H, Zhou H and Hickford JGH. Diversity of the glycine/tyrosine-rich keratin-associated protein 6 gene (KAP6) family in sheep. Molecular Biology Reports 2011, 38:31-5".

The PCR primers designed for *KRTAP6* worked well and strong amplicons were obtained for all of the sheep blood samples under the conditions described. These amplicons exhibited polymorphism upon SSCP analysis and in total five unique banding patterns could be detected (Figure 2-3). Sequencing of the amplicons revealed that these five unique patterns represented five different nucleotide sequences (*A* to *E*). Between two to four different sequences were identified in individual sheep (Figure 2-3).

Of the five sequences identified, sequence *D* was identical to a published ovine *KRTAP6-1* sequence (GenBank M95719), whereas the remaining sequences were novel, but the closest sequence homology was with the *KRTAP6* sequences from human, sheep, goats and cattle. This suggests that the sequences obtained here represent ovine *KRTAP6* sequences. These sequences have been deposited into the GenBank with accession numbers GU319872 - GU319876.

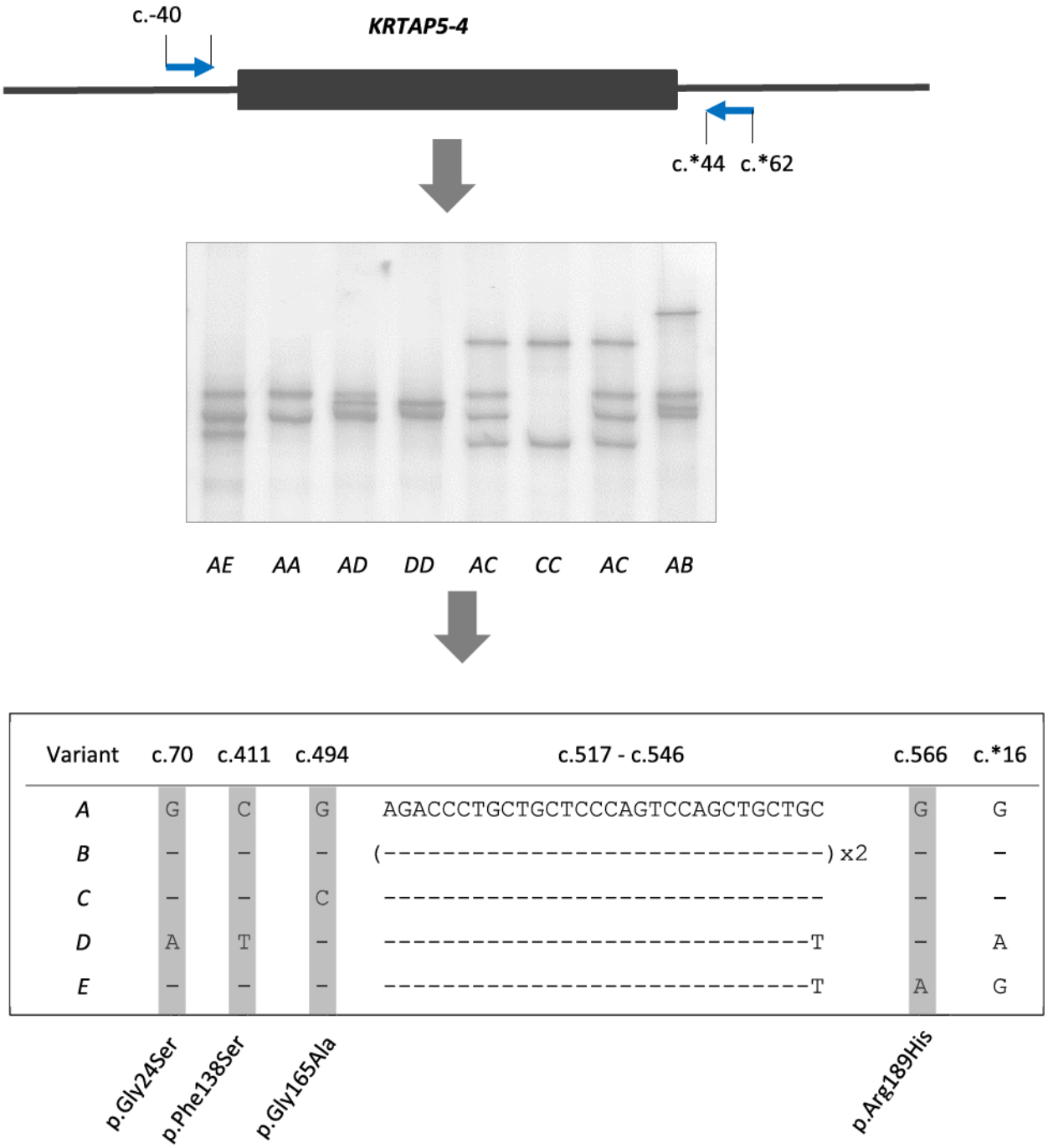
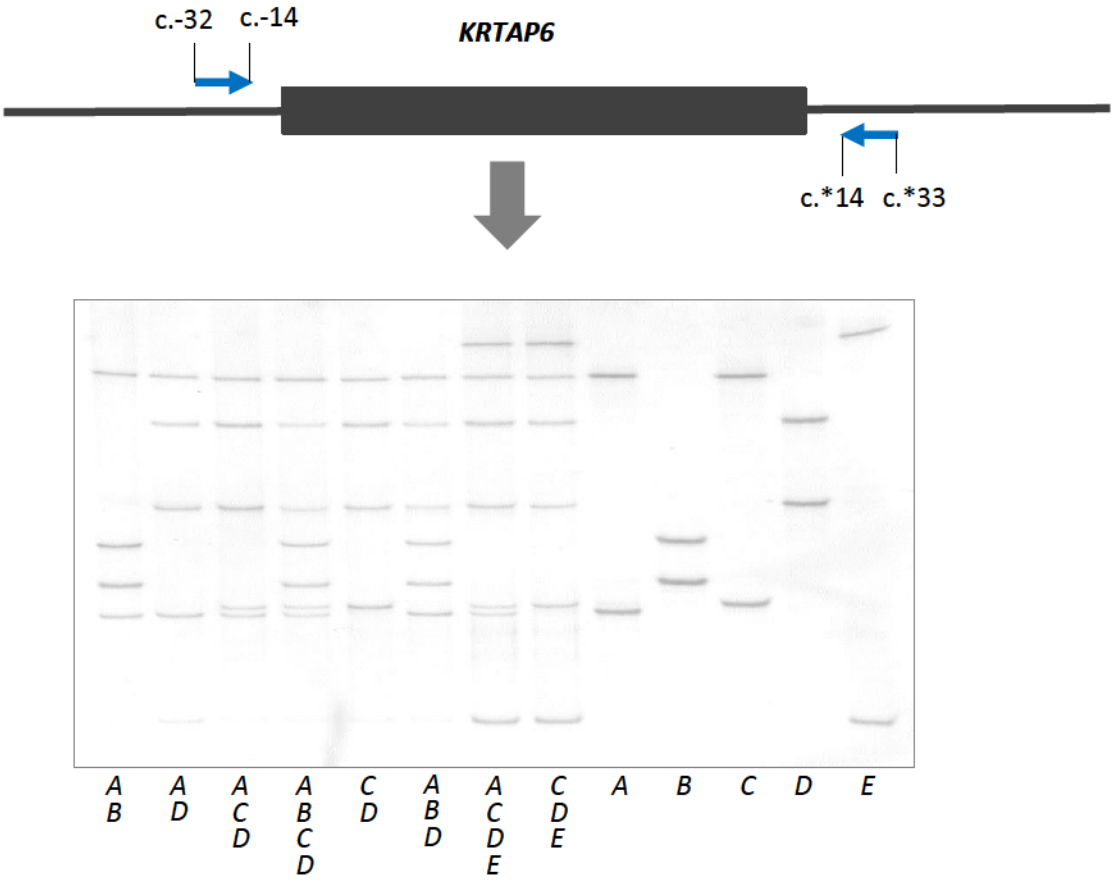


Figure 2-2. Variation identified in ovine *KRTAP5-4*. The entire coding sequence of *KRTAP5-4* was amplified by PCR using primers KRTAP5-4-up and KRTAP5-4-dn with the location of primers indicated. Five unique PCR-SSCP banding patterns representing five allelic variants (*A* to *E*) in either homozygous or heterozygous genotypes are shown. The non-synonymous SNPs are shaded with the corresponding amino acid substitutions being indicated. The 30 bp repeat sequence encoding for a cysteine-rich decapeptide sequences "RPCCSQSSCC" is shown. Nucleotide positions refer to GenBank X73434 and the numbering of nucleotide and amino acids follow the HGVS nomenclature (den Dunnen & Antonarakis, 2000).

(A)



(B)

```
KAP6-1      MCGYYGNYYGGLGCGSYSYGGLGCGYGSCYGSGFRRLGCGYG      42
KAP6*D      ------------------------------------------      42
KAP6*B      -------------------------..............      27
KAP6-2      .......GG-Y------G-------------NY---------      35
KAP6*A      -----------------G-----------------------      42
KAP6*C      -----------------G-----------------------      42
bKAP6-3     -----------------G-----------------------      42
KAP6*E      -----------------G-----Y---CGS------------      42


KAP6-1      CGYGYGSRSL.CGSGYGYGSRSLCGSGYGCGSGYGSGFGYYY      83
KAP6*D      ----------.------------------------------      83
KAP6*B      ....------.------------------------------      64
KAP6-2      ----------.------C---P-Y-C---------------.      75
KAP6*A      ----------.------C---P-Y-C----------------      83
KAP6*C      ----------.------C---P-Y-C----------------      83
bKAP6-3     SC--S-F-R-G--Y-C---YG-RSLC-C-Y-C---------C      84
KAP6*E      SC--S-F-R-G--Y-C--..........-Y-C----------      74
```

Figure 2-3. Diversity of ovine *KRTAP6*. (A) Five unique PCR-SSCP patterns, representing five different DNA sequences (*A* to *E*) were observed in sheep (lanes 1 to 8), with the patterns of these individual DNA sequences excised from the polyacrylamide gel and re-amplified being shown in lanes 9 to 13. (B) Sequence comparison revealed that these five sequences could be divided into three different groups as shown in different backgrounds. All amino acid sequences are deduced from DNA sequences except for KAP6-2 which was a partial sequence derived by protein sequencing (Gillespie, 1990). Amino acids identical to the top sequence are present as dashes, and dots have been introduced to improve the alignment. GenBank accession number for KAP6-1 is M95719.

Sequence comparison revealed that these five sequences could be divided into three groups (Figure 2-3). Sequences *B* and *D* were very similar to each other, except for a 57 bp deletion or insertion in the coding region and one SNP downstream of the coding region, but different to the other *KRTAP6* sequences. These two sequences were closely related to the published ovine *KRTAP6-1* sequence (GenBank M95719) and formed one group.

Sequences *A* and *C* were highly homologous (99.7%) to each other, but different to the other *KRTAP6* sequences identified and they appeared to form another group. There was only a single synonymous nucleotide difference between the *A* and *C* sequences, but these two sequences showed 10 and 11 nucleotide differences respectively to sequence *D*, the next most similar sequence among the other *KRTAP6* sequences identified. This variation results in five amino acid substitutions for sequences *A* and *C*, compared to *D*.

The remaining sequence isolated in this study (sequence *E*) did not appear to have homology with other *KRTAP*6 sequences and formed another group by itself. At the predicted amino acid sequence level, there are at least 25 amino acid differences between sequence *E* and other *KRTAP6* sequences. However, the closest sequence similarity was found with the putative *KRTAP6-3* sequence from cattle (GenBank XM_883346). There were only five amino acid differences between the putative polypeptide produced by ovine *KRTAP6*E* and that produced from the bovine *KRTAP6-3* sequence, with the exception of a ten amino acid deletion/insertion. The *E* sequence was only detected in 11% of the sheep investigated.

Two allele variants observed for ovine *KRTAP7-1*

This work has been published in "Gong H, Zhou H, Plowman JE, Dyer JM and Hickford JGH. Search for variation in the ovine KAP7-1 and KAP8-1 genes using polymerase chain reaction-single-stranded conformational polymorphism screening. DNA and Cell Biology 2012, 31:367-70".

The PCR primers designed for ovine *KRTAP7-1* produced PCR amplicons of the expected size (approximately 327 bp) with all the sheep DNA samples. These amplicons exhibited different banding patterns upon SSCP analysis and two unique banding patterns could be detected (Figure 2-4). Sequencing of the amplicons revealed that these two unique patterns represented DNA sequences that had a single nucleotide difference (c.173G/A). One of the sequences was identical to the published ovine *KRTAP7-1* sequence (GenBank X05638) and named *KRTAP7-1*A*, whereas the other was novel and named *KRTAP7-1*B*. The nucleotide substitution detected in ovine *KRTAP7-1* was non-synonymous and would result in an amino acid change (p.Ser58Asn) if the gene was to be expressed.

These allelic sequences have been deposited into the GenBank with accession numbers JN091630 and JN091631. The frequencies of *KRTAP7-1*A* and *KRTAP7-1*B* in the sheep investigated were 77% and 23%, respectively.

Five allele variants detected in ovine *KRTAP8-1*

This work has been published in "Gong H, Zhou H, Plowman JE, Dyer JM and Hickford JGH. Search for variation in the ovine KAP7-1 and KAP8-1 genes using polymerase chain reaction-single-stranded conformational polymorphism screening. DNA and Cell Biology 2012, 31:367-70".

PCR amplification using the *KRTAP8-1* primers generated amplicons of the expected size (approximately 258 bp) with all the sheep DNA samples. These amplicons displayed different banding patterns on SSCP gels and five unique banding patterns were detected (Figure 2-5). One or a combination of two different banding patterns was observed for each individual sheep, which is consistent with them being either homozygous or heterozygous at this locus. Sequencing of the amplicons revealed five different DNA sequences. All of the sequences were highly similar to, but none of them was identical, to the published ovine *KRTAP8-1* sequence (GenBank X05639) (Figure 2-5). These sequences were named *KRTAP8-1*A* to *KRTAP8-1*E*, and they were deposited into the GenBank with the accession numbers JN091632 - JN091636. Alleles *A* to *E* were detected at frequencies of 68.3%, 13.3%, 3.3%, 10.0%, and 5.0%, respectively, in the sheep investigated.

Four SNPs were identified in ovine *KRTAP8-1* in these five alleles and three of them were located in the coding region. Of these, only one substitution was non-synonymous and it would result in an amino acid change (p.Tyr34Asn).

It is notable that all five *KRTAP8* sequences have different nucleotides to the GenBank sequence (X05639) at positions c.48 (T>C) and c.152 (C>T), the latter causing a Ser/Phe amino acid substitution.

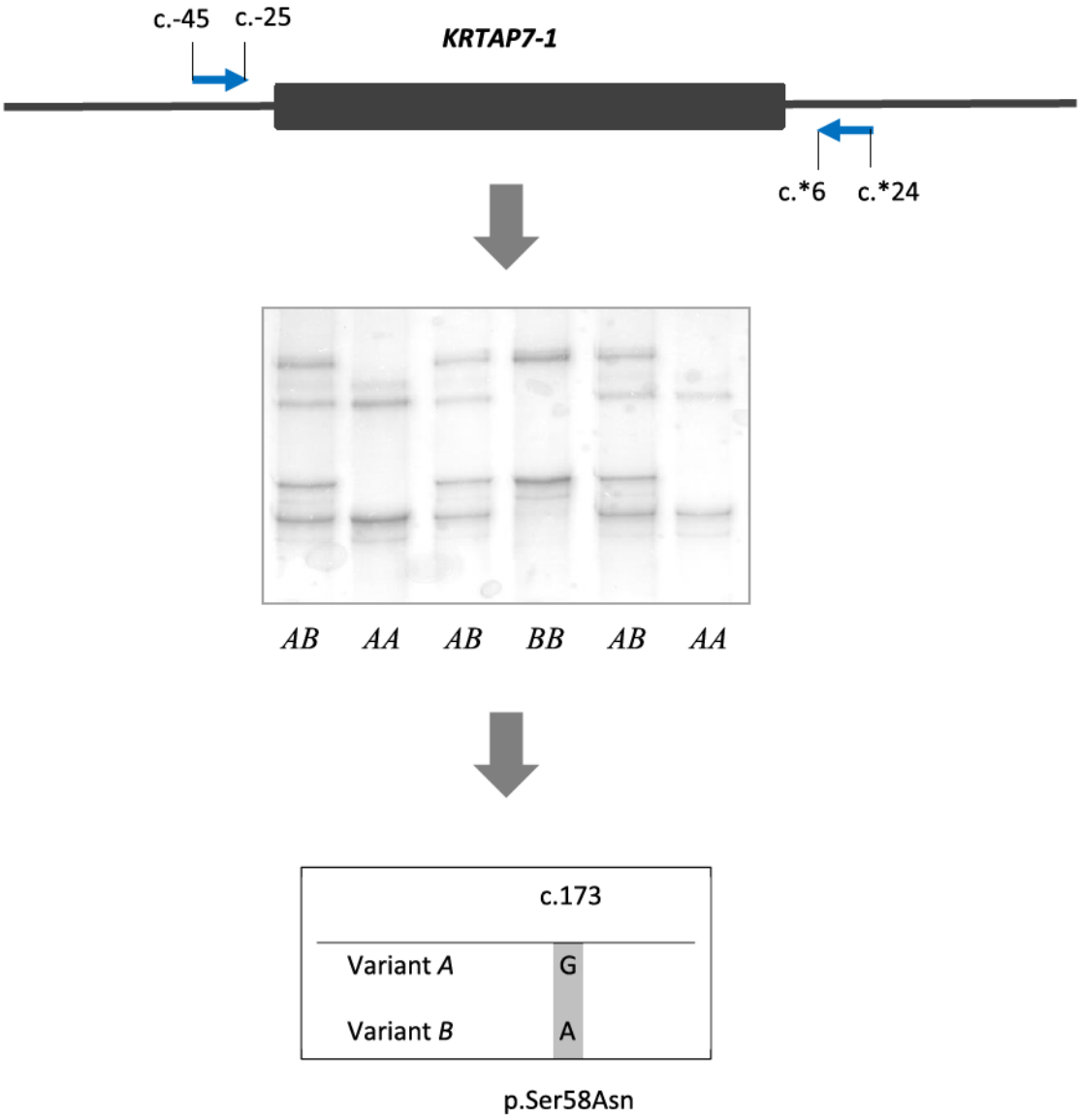
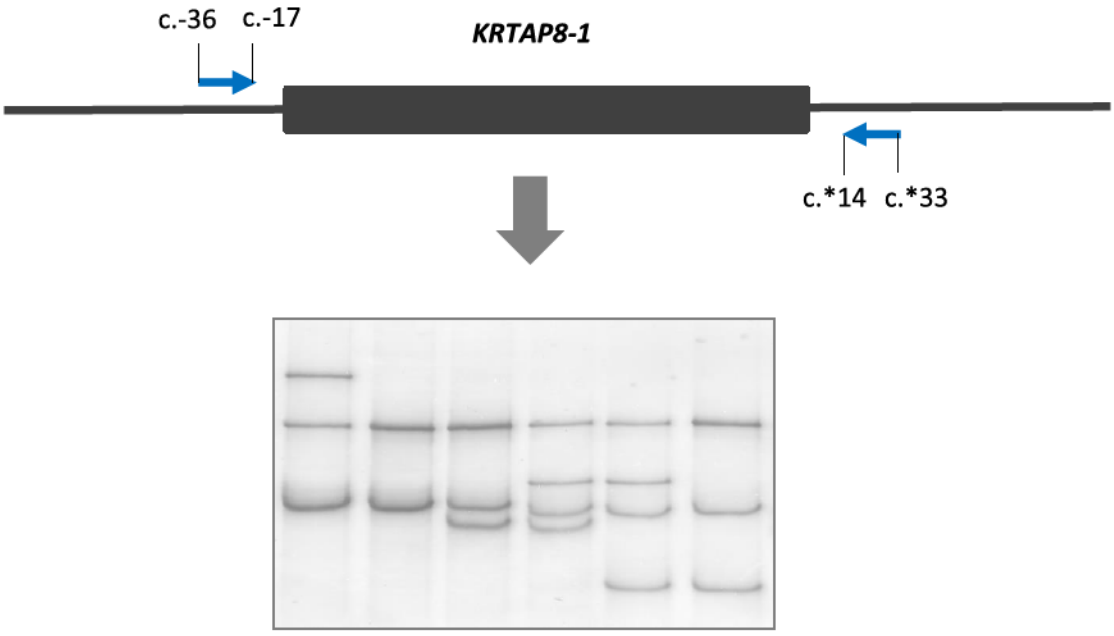


Figure 2-4. Variation identified in ovine *KRTAP7-1*. The entire coding sequence of *KRTAP7-1* was amplified using primers KRTAP7-up and KRTAP7-dn with the location of the primers indicated. Two unique PCR-SSCP banding patterns representing two allelic variants (*A* and *B*) in both homozygous and heterozygous genotypes are shown. The non-synonymous SNP is shaded and the corresponding amino acid change is indicated. Nucleotide positions refer to GenBank X05638 and the numbering of nucleotide and amino acids follow the HGVS nomenclature (den Dunnen & Antonarakis, 2000).



	c.-17	c.39	c.48	c.100	c.117	c.152
Variant *A*	C	C	C	T	T	T
Variant *B*	T	-	-	-	-	-
Variant *C*	-	T	-	-	-	-
Variant *D*	-	-	-	A	-	-
Variant *E*	-	-	-	-	C	-
			#	p.Tyr34Asn		#

Figure 2-5. Variation identified in ovine *KRTAP8-1*. The entire coding sequence of *KRTAP8-1* was amplified using primers KRTAP8-up and KRTAP8-dn with the location of the primers indicated. Five unique PCR-SSCP banding patterns representing five allelic variants (*A* to *E*) in either homozygous or heterozygous genotypes are shown. The non-synonymous SNP identified among these variants is shaded with the corresponding amino acid substitution being shown. The nucleotides that are different to GenBank X05639 are indicated by #. Nucleotide positions refer to GenBank X05639 and the numbering of nucleotide and amino acids follow the HGVS nomenclature (den Dunnen & Antonarakis, 2000).

2.3.2 Characterization of New Ovine KAP Genes

Identification of ovine *KRTAP1-2* with nine allelic variants

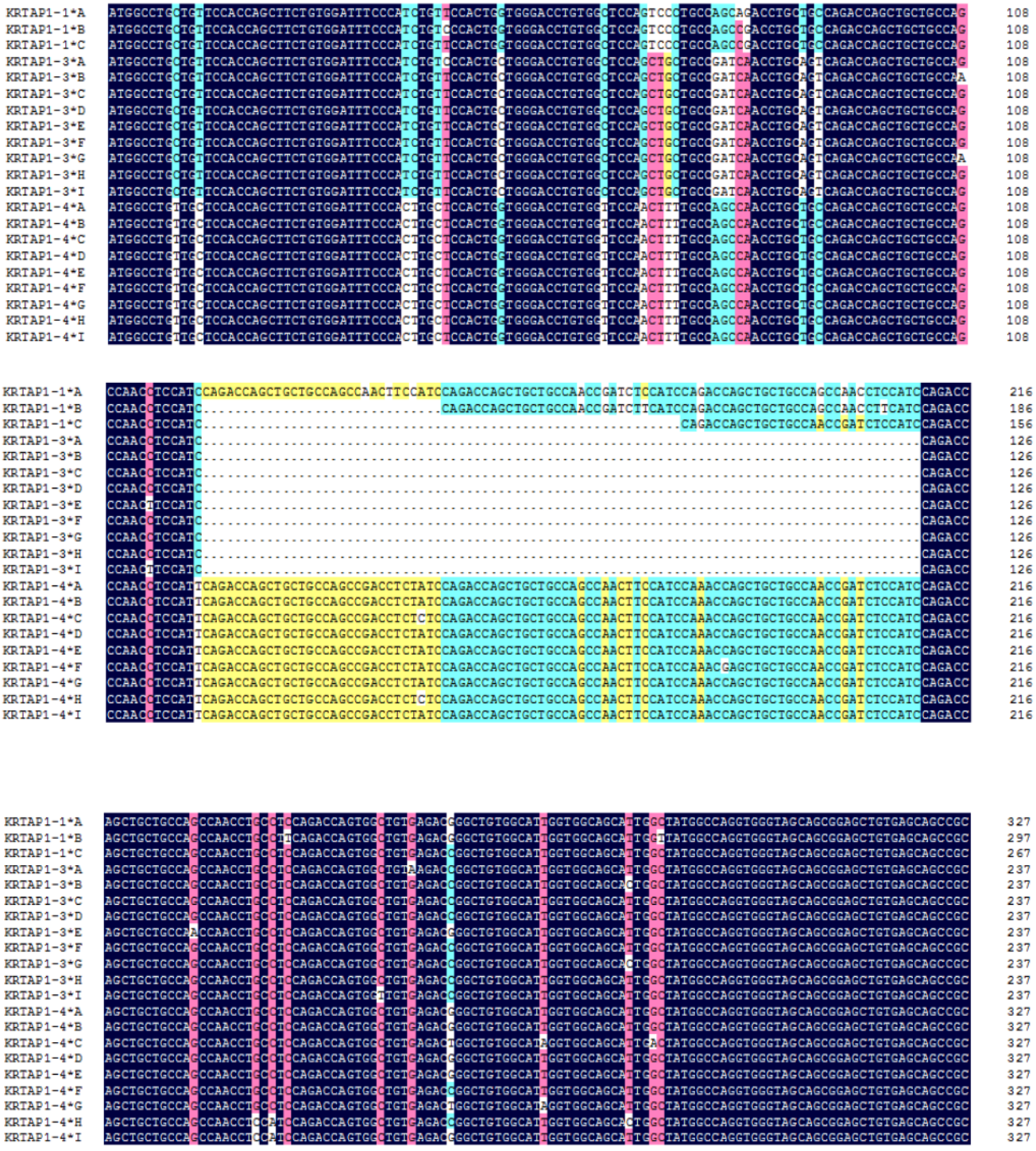
This work has been published in "Gong H, Zhou H, Yu Z, Dyer J, Plowman JE and Hickford JGH. Identification of the ovine keratin-associated protein KAP1-2 gene (KRTAP1-2). Experimental Dermatology 2011, 20: 815-9".

Alignment of all known ovine *KRTAP1-1*, *KRTAP1-3* and *KRTAP1-4* sequences revealed a 87 bp sequence that was conserved across these genes and that would notionally translate to a 29 amino acid sequence (PSCCQLYYAQASCCRPSYCGQSCCRPACC) found in the ovine KAP1-2 protein sequence of KAP1-2 (Swiss-Prot P02439.1) (Figure 2-6).

A BLAST search of the Ovine Genome Sequence Assembly v1.0 using this 87 bp nucleotide sequence revealed two homologous regions in sheep chromosome 11. The first region was OAR11:43296072_43302531 which matched at the sequence level with a genomic clone sequence (Powell *et al.*, 1983) covering *KRTAP1-4* and *KRTAP1-1* (X01610). The second region was OAR11:43306563_43307552, located approximately 4 kb downstream of the first region (Figure 2-7). The partial sequence retrieved from the second region did not match any known ovine gene, suggesting it was a newly identified gene. Approximately 6 kb downstream of the second region, there was a region (OAR11:43313559_43315018) containing a sequence matched with the previously reported ovine *KRTAP1-3* sequence (X02925) (Powell *et al.*, 1983) (Figure 2-7).

Two PCR primers designed based on the sequences of the second region produced an amplicon of 909 bp from sheep genomic DNA. The sequence of this amplicon matched with the sequence of the second region, although only a partial gene sequence was available from the Ovine Genome Sequence Assembly v1.0. This suggests that the PCR primers amplified the targeted gene and the amplicon obtained was derived from the OAR11:43306563_43307552 region.

The 909 bp amplicon contained an open reading frame of 474 bp that was predicted upon translation to encode a polypeptide of 157 amino acids. The sequence of this polypeptide was identical to the published ovine KAP1-2 protein sequence (Swiss-Prot P02439.1), with the exception of two amino acids [p.29Ser and p.30Cys predicted in this study *versus* p.29Cys and p.30Ser reported previously by Elleman & Dopheide (1972)] (Figure 2-8). This suggested that the sequence represented the ovine KAP1-2 gene. The sequence has been entered into GenBank with accession number HQ897973.


KRTAP1-1*A
KRTAP1-1*B
KRTAP1-1*C
KRTAP1-3*A
KRTAP1-3*B
KRTAP1-3*C
KRTAP1-3*D
KRTAP1-3*E
KRTAP1-3*F
KRTAP1-3*G
KRTAP1-3*H
KRTAP1-3*I
KRTAP1-4*A
KRTAP1-4*B
KRTAP1-4*C
KRTAP1-4*D
KRTAP1-4*E
KRTAP1-4*F
KRTAP1-4*G
KRTAP1-4*H
KRTAP1-4*I

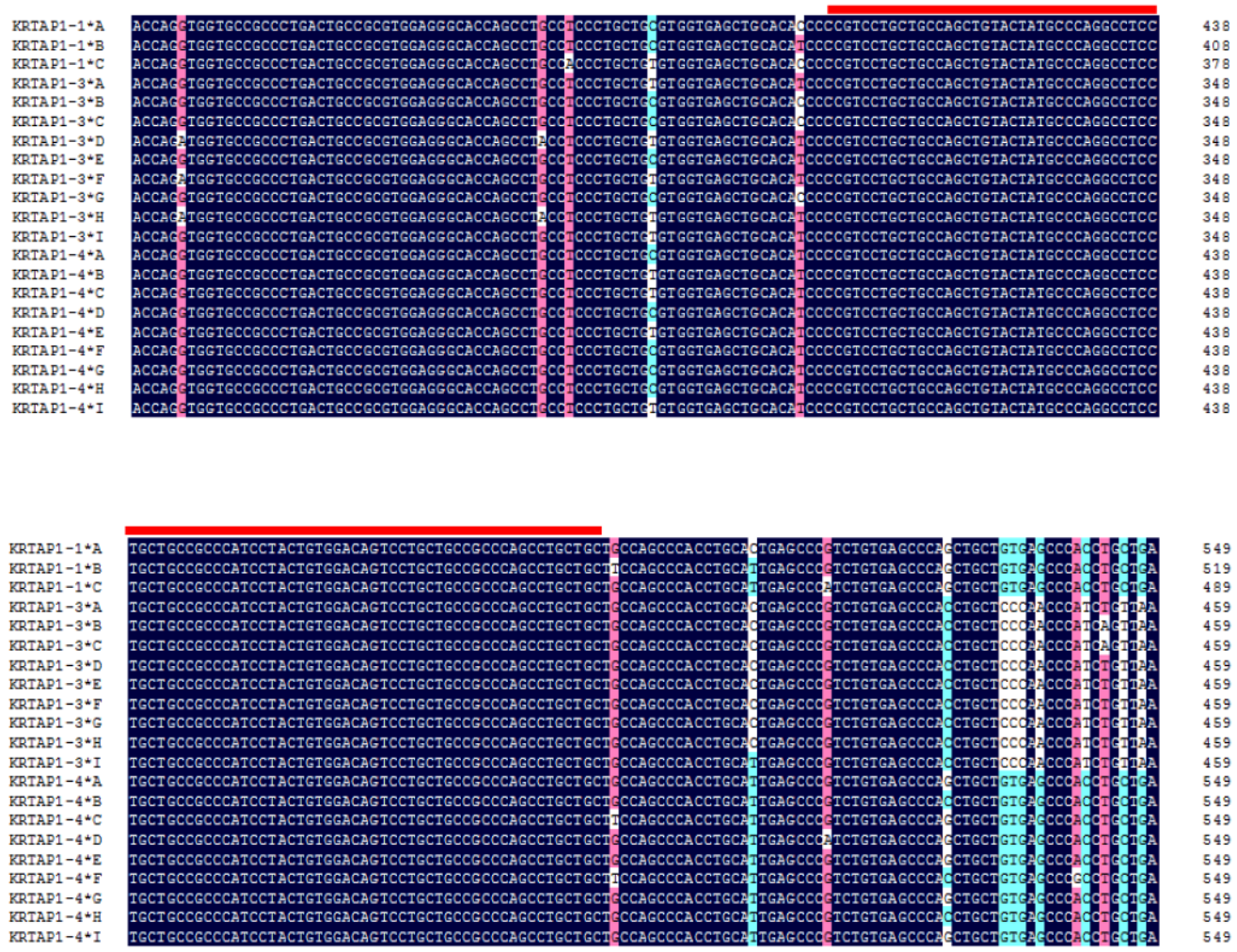


Figure 2-6. Alignment of the ovine *KRTAP1-1*, *KRTAP1-3* and *KRTAP1-4* genes. A 87 bp sequence that was conserved across all known *KRTAP1-1*, *KRTAP1-3* and *KRTAP1-4* sequences and that would notionally translate to a sequence found in the ovine KAP1-2 protein sequence (Swiss-Prot P02439.1) is indicated by the red bar above the sequences. The GenBank accession numbers for the sequences aligned are: L33885 to L33887 for KAP1-1*A to KAP1-1*C, AY835589 to AY835597 for KAP1-3*A to KAP1-3*I, and GQ507741 to GQ507749 for KAP1-4*A to KAP1-4*I, respectively.

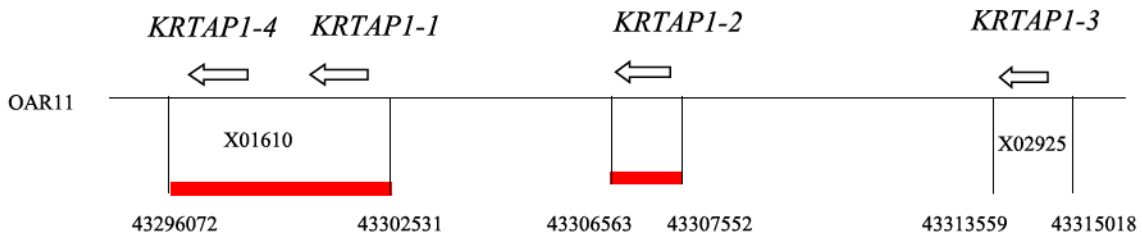


Figure 2-7. Location of four putative KAP1-n genes on sheep chromosome 11. Regions homologous to the conserved 87 bp nucleotide sequence are indicated by red bars. Arrows indicate the direction of transcription. Nucleotide positions refer to the Sheep Genome Assembly v1.0. GenBank accession numbers for *KRTAP1-4, KRTAP1-1* and *KRTAP1-3* are shown. The failure to detect an 87 bp homologous region in *KRTAP1-3* is probably due to the existing sequence being partial.

Within the coding region, the KAP1-2 gene exhibited a high sequence homology to KAP1-1, KAP1-3 and KAP1-4 genes and a number of conserved sequence motifs could be identified among the polypeptides encoded by these genes (Figure 2-8). Despite the overall conservation of the ovine KAP1-n family members, each member possesses a number of unique sequence motifs. Comparing with other KAP1-n members, the notional KAP1-2 polypeptide from the sequence identified had four unique amino acid residues (p.16Ser, p.17Val, p.25Gly and p.29Ser) and also lacked five residues at the C-terminus (Figure 2-8). However, the nucleotide sequences were less well conserved outside the coding region, with only two small regions being comparable and some additional homology through the putative TATA box. Beyond these similarities, the remaining sequences were markedly different (Figure 2-9).

PCR amplification using the two internal PCR primers (KRTAP1-2s-up and KRTAP1-2s-dn) produced amplicons of the expected size (approximately 557 bp) from all sheep blood samples. These amplicons exhibited different banding patterns upon SSCP analysis and nine unique banding patterns could be identified (Figure 2-10). Either one or a combination of two patterns was observed for each sheep.

Sequencing of PCR amplicons representative of these unique SSCP patterns revealed nine different DNA sequences. One of these sequences was identical to the original 909 bp *KRTAP1-2* sequence described above (GenBank HQ897973), whereas the remaining eight sequences were unique but showed high similarity to the 909 bp *KRTAP1-2* sequence (GenBank HQ897973). These sequences were named alleles *A-I* and deposited into the GenBank with the accession numbers HQ897974 - HQ897982.

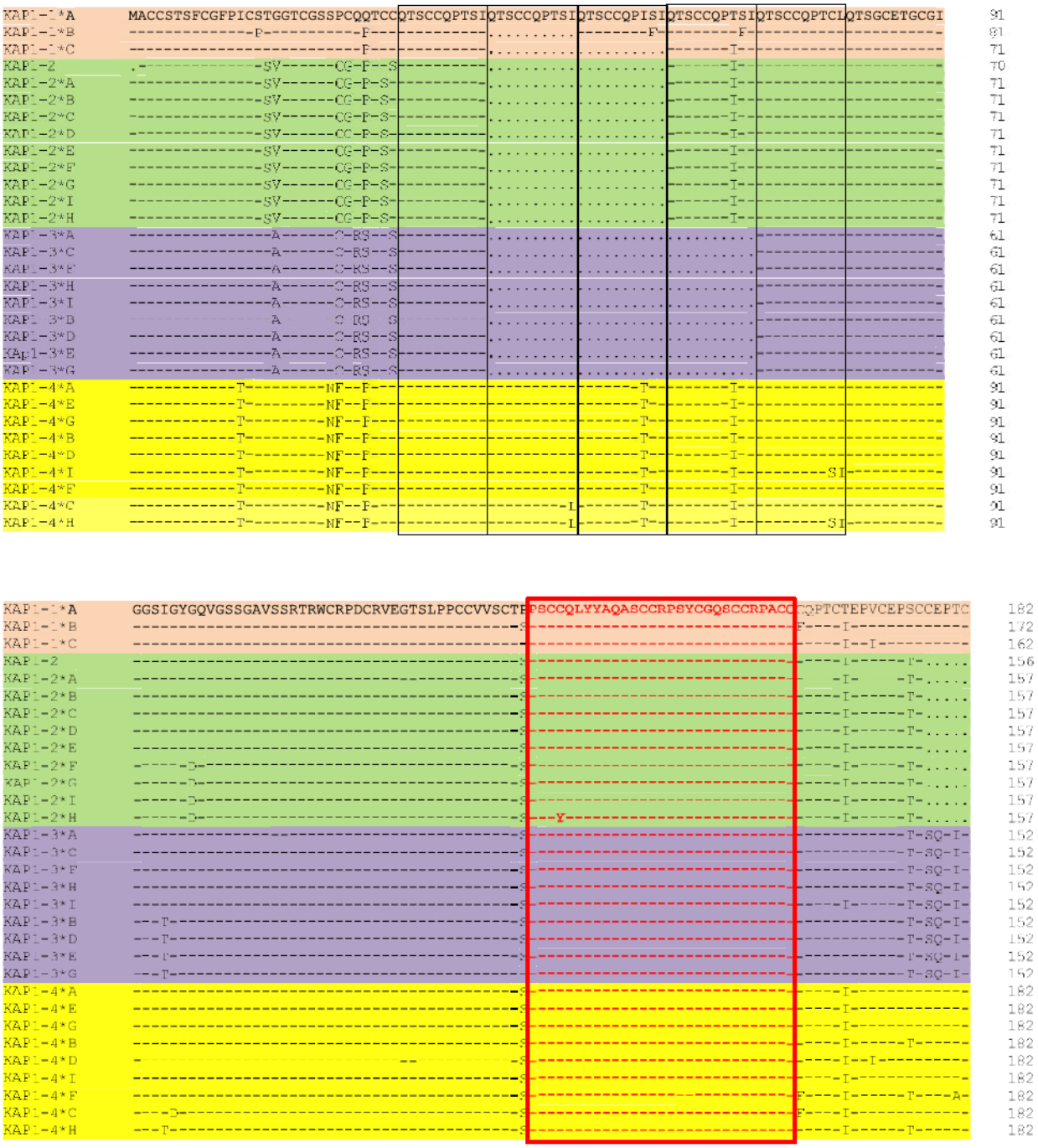


Figure 2-8. Alignment of the predicted amino acid sequences for various alleles of KAP1-n genes, together with the KAP1-2 protein sequence (Swiss-Prot P02439.1). Amino acids identical to the top sequence are presented by dashes and dots have been introduced to improve the alignment. The decapeptide repeats "QTSCCQPXXX" are shown in black boxes. The 29 amino acid sequence encoded by the 87 bp conserved sequence used to BLAST the ovine genome sequence is bolded and shown in a red box. Sequence variants of individual KAP1-n members are grouped and shaded in different colours.

There were 10 single nucleotide substitutions identified in this gene (Figure 2-10). Eight of the substitutions were located in the coding region. Of these eight substitutions, three were non-synonymous. These were c.233G/A, c.353G/A, and c.467C/G, and they would result in amino acid changes of p.Gly78Asp, p.Cys118Tyr and p.Thr156Ser, respectively (Figure 2-8 and Figure 2-10).

a)

b)

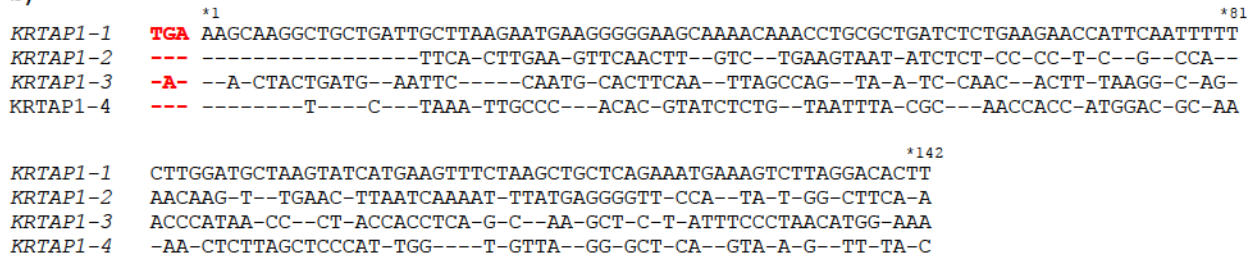
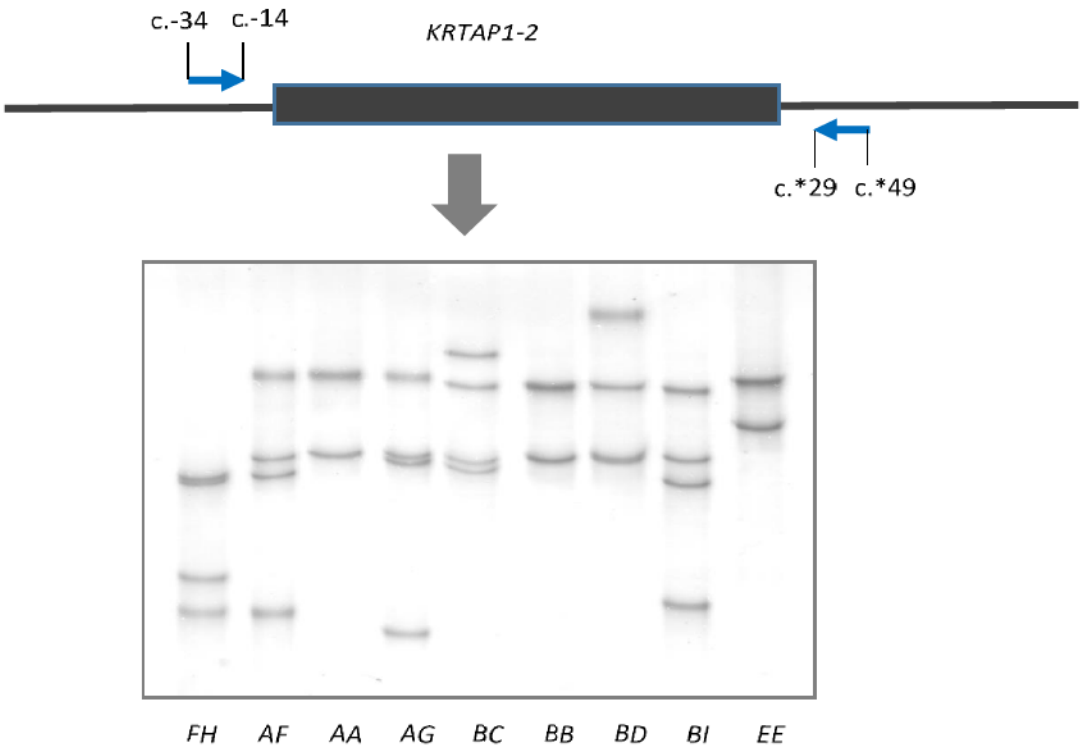


Figure 2-9. Alignment of the 5' (a) and 3' (b) flanking sequences of the ovine KAP1-n genes. Nucleotides identical to the top sequence are presented by dashes and dots have been introduced to improve the alignment. Three regions that appear to be conserved among these genes are shaded. The notional TATA sequence is in bold and indicated by a horizontal bar. The putative cap-site nucleotides for KAP1-1, KAP1-3 and KAP1-4 are boxed, while the position in KAP1-2 cannot be precisely determined in this study. The initiation and stop codons are bolded and shown in red. The sequences of *KRTAP1-1* and *KRTAP1-4* were obtained from GenBank X01610 and that of *KRTAP1-3* is from GenBank X02925. Nucleotide positions in the 5' flanking region have a "-" prefix and refer to the number of nucleotides upstream of the initiation codon. Nucleotide positions in the 3' flanking region have a "*" prefix and refer to the number of nucleotides downstream of the stop codon.



Variant	c.-1	c.99	c.201	c.213	c.233	c.249	c.324	c.353	c.467	c.*5
A	C	T	C	T	G	C	T	G	C	A
B	-	-	G	-	-	-	-	-	-	-
C	-	-	A	A	-	-	-	-	-	G
D	-	-	-	A	-	-	-	-	-	-
E	-	-	G	-	-	T	C	-	G	-
F	T	-	-	-	A	-	-	-	-	-
G	-	C	-	-	A	T	-	-	-	-
H	-	-	A	-	A	T	C	A	-	-
I	-	-	A	-	A	-	-	-	-	-

p.Gly78Asp p.Cys118Tyr p.Thr156Ser

Figure 2-10. Variation identified in ovine *KRTAP1-2*. The entire coding sequence of *KRTAP1-2* was amplified using primers KRTAP1-2s-up and KRTAP1-2s-dn with the location of primers indicated. Nine unique PCR-SSCP banding patterns representing nine allelic variants (*A* to *I*) in either homozygous or heterozygous genotypes are shown. Ten SNPs were identified and the non-synonymous SNPs are shaded with the corresponding amino acid substitutions being shown. Nucleotide positions refer to GenBank HQ897973 and the numbering of nucleotide and amino acids follow the HGVS nomenclature (den Dunnen & Antonarakis, 2000).

Identification of ovine *KRTAP8-2* with two allelic variants

This work has been published in "Gong H, Zhou H, Dyer J and Hickford JGH. The sheep KAP 8-2 gene a new KAP8 family member that is absent in humans. SpringerPlus 2014, 3:528".

A BLAST search of the Ovine Genome Assembly v2.0 using the caprine *KRTAP8-2* coding sequence (AY510123) revealed a region on sheep chromosome 1 (OAR1:123005473_123005664; E = e-101) that contained a 192 bp open reading frame and that had 99% identity with the caprine gene. Near this region, seven previously described ovine KAP genes were also identified and these (including *KRTAP8-2*) were *KRTAP11-1, KRTAP7-1, KRTAP8-1, KRTAP8-2, KRTAP6-2, KRTAP6-1, KRTAP13-3* and *KRTAP24-1* (in order from the centromere) (Figure 2-11). The open reading frame identified had high homology with sheep skin expressed sequence tags (ESTs) in GenBank and identical sequences covering the entire open reading frame were found in 44 EST sequences derived from skin tissues (Table 2-2).

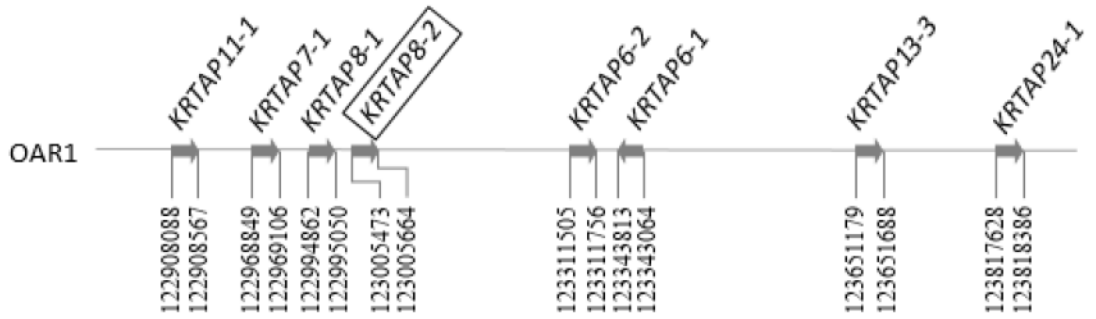


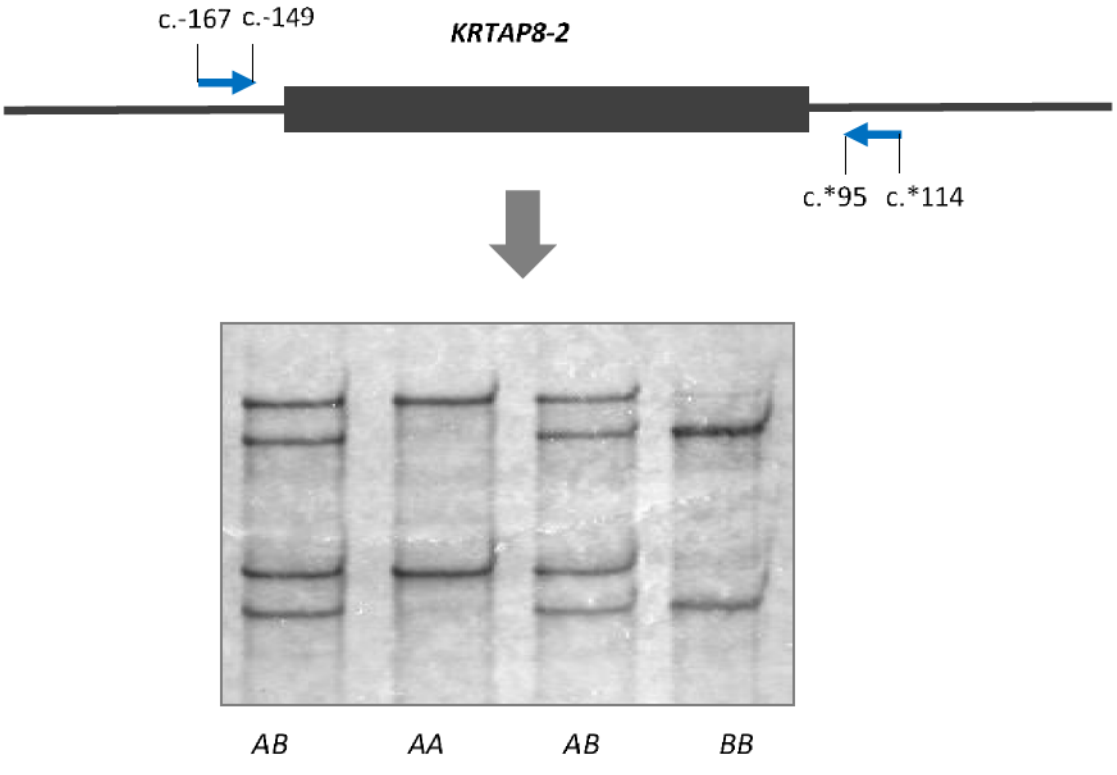
Figure 2-11. Location of the putative *KRTAP8-2* (boxed) together with seven other *KRTAPs* on sheep chromosome 1. The coding regions of individual *KRTAPs* are shown, with the nucleotide positions referring to Ovine Genome Assembly v2.0. Arrows represent the direction of transcriptions.

PCR amplification of the entire open reading frame and its flanking sequence generated amplicons with the expected size of 473 bp. SSCP analysis of these amplicons revealed two unique banding patterns, with either one or a combination of two patterns being observed in each sheep (Figure 2-12).

Sequencing of PCR amplicons revealed that these two PCR-SSCP patterns represented two different DNA sequences. These sequences differed from each other by one nucleotide, 21 bp upstream of the TATA box. Neither of the sequences was identical to the sequence reported in v2.0 of the Ovine Genome Assembly, with three nucleotide differences being detected in the 3'-UTR. This likely reflects either additional genetic variation in the gene, or sequencing/assembly errors within v2.0.

Table 2-2. Sheep skin ESTs identical to the ovine KAP8-2 gene

GenBank accession no.	Score (bit)	Expect	Identity
O781125.1	355	5.00E-95	192/192 (100%)
GO778998.1	355	5.00E-95	192/192 (100%)
GO777073.1	355	5.00E-95	192/192 (100%)
GO780070.1	355	5.00E-95	192/192 (100%)
GO779393.1	355	5.00E-95	192/192 (100%)
GO766145.1	355	5.00E-95	192/192 (100%)
GO768646.1	355	5.00E-95	192/192 (100%)
GO739616.1	355	5.00E-95	192/192 (100%)
GO736506.1	355	5.00E-95	192/192 (100%)
GO738396.1	355	5.00E-95	192/192 (100%)
GO738173.1	355	5.00E-95	192/192 (100%)
GO710784.1	355	5.00E-95	192/192 (100%)
GO710584.1	355	5.00E-95	192/192 (100%)
GO709687.1	355	5.00E-95	192/192 (100%)
GO705655.1	355	5.00E-95	192/192 (100%)
GO707617.1	355	5.00E-95	192/192 (100%)
GO707056.1	355	5.00E-95	192/192 (100%)
GO706322.1	355	5.00E-95	192/192 (100%)
GO703222.1	355	5.00E-95	192/192 (100%)
GO694213.1	355	5.00E-95	192/192 (100%)
GO693222.1	355	5.00E-95	192/192 (100%)
GO693118.1	355	5.00E-95	192/192 (100%)
GO691792.1	355	5.00E-95	192/192 (100%)
GO691319.1	355	5.00E-95	192/192 (100%)
GO679286.1	355	5.00E-95	192/192 (100%)
GO678916.1	355	5.00E-95	192/192 (100%)
GO678591.1	355	5.00E-95	192/192 (100%)
EE852829.1	355	5.00E-95	192/192 (100%)
EE849147.1	355	5.00E-95	192/192 (100%)
EE848083.1	355	5.00E-95	192/192 (100%)
EE847151.1	355	5.00E-95	192/192 (100%)
EE758126.1	355	5.00E-95	192/192 (100%)
EE758032.1	355	5.00E-95	192/192 (100%)
EE757975.1	355	5.00E-95	192/192 (100%)
EE757004.1	355	5.00E-95	192/192 (100%)
EE756700.1	355	5.00E-95	192/192 (100%)
EE756552.1	355	5.00E-95	192/192 (100%)
EE756175.1	355	5.00E-95	192/192 (100%)
EE755582.1	355	5.00E-95	192/192 (100%)
EE753480.1	355	5.00E-95	192/192 (100%)
EE753151.1	355	5.00E-95	192/192 (100%)
EE751683.1	355	5.00E-95	192/192 (100%)
EE751546.1	355	5.00E-95	192/192 (100%)
EE751314.1	355	5.00E-95	192/192 (100%)



	c.-101
Variant *A*	C
Variant *B*	T

Figure 2-12. Variation identified in ovine *KRTAP8-2*. The entire coding sequence of *KRTAP8-2* was amplified using primers KRTAP8-2-up and KRTAP8-2-dn with the location of primers indicated. Two unique PCR-SSCP banding patterns representing two allelic variants (*A* and *B*) in both homozygous and heterozygous genotypes are shown. Nucleotide positions refer to GenBank KF220646 and the numbering of nucleotide and amino acids follow the HGVS nomenclature (den Dunnen & Antonarakis, 2000).

The sheep sequences identified here did not share significant homology with any other known ovine *KRTAP* sequence, but sequence similarity was found with the *KRTAP8-2* sequences from goats (AY510123) and reindeer (EF407854), with 99% and 95% similarity, respectively in the coding region. These sequences were assumed to represent allelic variants of ovine *KRTAP8-2* and were named variants *A* and *B*. They were placed in GenBank under the accession numbers KF220646 and KF220646, respectively.

Variant *A* was found most frequently (at a frequency of 90.5%), while variant *B* was less common (at a frequency of 9.5%) in the Romney-cross sheep investigated. Three genotypes were observed with frequencies of 83%, 15% and 2% for *AA, AB* and *BB*, respectively.

The putative *KRTAP8-2* sequence would encode a 63 amino acid polypeptide that contained a high level of glycine (23.8 mol%) and tyrosine (20.6 mol%), accounting for 44.4 mol% in total of the amino acid content. It had a moderate amount of phenylalanine (9.5 mol%) and serine (7.9 mol%), but a relatively low cysteine content (3.2 mol%). This polypeptide also possessed 3.2 mol% aspartic acid and 1.6 mol% glutamic acid, amino acids that are absent in other HGT-KAPs. The calculated isoelectric point (pI) of the protein was 6.3 (Table 2-3).

Identification of ovine *KRTAP11-1* with six allelic variants

This work has been published in "Gong H, Zhou H, Dyer JM and Hickford JGH. Identification of the ovine KAP11-1 gene (KRTAP11-1) and genetic variation in its coding sequence. Molecular Biology Reports 2011, 38: 5429-33".

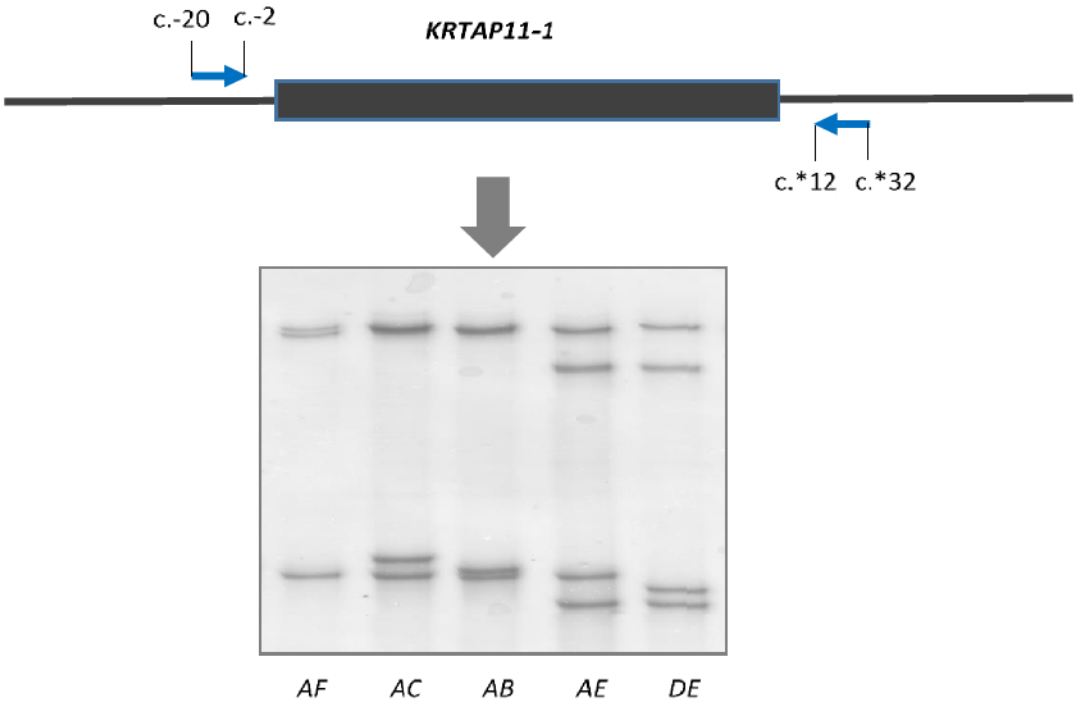
PCR amplification using the *KRTAP11-1* primers generated a PCR amplicon with the expected size (532 bp) from all sheep blood samples. These amplicons exhibited different banding patterns upon SSCP analysis and six unique patterns could be identified (Figure 2-13). Either one or a combination of two different SSCP patterns was observed for each amplicon (Figure 2-13).

Sequencing of PCR amplicons representative of these unique SSCP patterns revealed six different nucleotide sequences. All of these sequences were different to the known ovine *KRTAP* sequences with a low sequence similarity being observed. However, at the predicted amino acid level, the closest homologues of these sequences were the *KRTAP11-1* sequences from cattle, human and mouse. The ovine sequences were named *KRTAP11-1* alleles *A-F* and deposited into GenBank with the accession numbers HQ595347 - HQ595352.

Table 2-3. Comparison of the amino acid content (mol%) and isoelectric point (pI) value of ovine KAP8-2 and other ovine HGT-KAPs

HGT-KAP*	Glycine	Tyrosine	Cysteine	Serine	Phenylalanine	Proline	Aspartic acid	Glutamic acid	pI
KAP6-1	37.4-37.5	21.7-23.4	9.4-10.8	14.5-15.6	1.6-2.4	0	0	0	8.1-8.3
KAP6-2	38.6	21.7	12.1	10.8	2.4	1.2	0	0	8.2
KAP7-1	22.4	11.8	5.9	12.9-14.1	10.6	7.1	0	0	8.7
KAP8-1	22.6	16.1-17.7	6.5	12.9	9.7	6.5	0	0	8.3
KAP8-2	23.8	20.6	3.2	7.9	9.5	6.4	3.2	1.6	6.3

* Sequences accession numbers: KAP6-1 (GU319873 and GU319875); KAP6-2 (GU3198772 and GU319874); KAP7-1 (JN091630 and JN091631); KAP8-1 (JN091632 to JN091636); and KAP8-2 (KF220646 and KF220646).



	c.69	c.93	c.285	c.331	c.423
Variant *A*	G	C	C	A	C
Variant *B*	-	T	-	-	-
Variant *C*	T	T	-	-	-
Variant *D*	-	-	-	-	T
Variant *E*	-	T	G	G	-
Variant *F*	-	-	A	-	-

p.Ser111Gly

Figure 2-13. Variation identified in ovine *KRTAP11-1*. The entire coding sequence of *KRTAP11-1* was amplified using primers KRTAP11-1-up and KRTAP11-1-dn with the location of primers indicated. Six unique PCR-SSCP banding patterns representing six allelic variants (*A* to *F*) in either homozygous or heterozygous genotypes are shown. The non-synonymous SNP identified in this study is shaded with the corresponding amino acid substitution being shown. Nucleotide positions refer GenBank HQ595347 and the numbering of nucleotide and amino acids follow the HGVS nomenclature (den Dunnen & Antonarakis, 2000).

Alleles *A* and *B* were the most common with allele frequencies of 42.1% and 36.7%, respectively. The next most common allele was *E* with a frequency of 10.4%, followed by allele *C* (with a frequency of 7.9%). Alleles *D* and *F* were the least common alleles with frequencies of 1.9% and 1.0%, respectively.

All the *KRTAP11-1* sequences had an open reading frame of 477 nucleotides, encoding a putative polypeptide of 159 amino acids. The polypeptide had a low content of alanine, histidine, lysine, methionine, asparagine and tryptophan (which only occurred once in the putative polypeptide). However, the putative polypeptide contained a moderate content of glycine (5.7 mol%), and a high level of serine, cysteine and threonine. These were the most common amino acids accounting for 18.2 - 18.9, 12.6 and 12.0 mol%, respectively. Approximately 20 serine and threonine resides were predicted to be phosphorylated using the NetPhos 2.0 Server (www.cbs.dtu.dk/services/NetPhos/) (Figure 2-14).

A total of five nucleotide substitutions were identified among the *KRTAP11-1* sequences, and all of these were located in the coding region (Figure 2-13). However, there was only one substitution which would result in an amino acid change (p.Ser111Gly) in the putative polypeptide (Figure 2-13 and Figure 2-14).

Identification of ovine *KRTAP13-3* with five allelic variants

This work has been published in "Gong H, Zhou H, Dyer JM, Plowman JE and Hickford JGH. Identification of the keratin-associated protein 13-3 (KAP13-3) gene in sheep. Open Journal of Genetics 2012, 1:60-4".

A PCR amplicon with the expected size (594 bp) was obtained from all sheep blood samples using the *KRTAP13-3* primers. The amplicons exhibited polymorphism upon SSCP analysis with five different SSCP patterns being detected (Figure 2-15).

Sequencing of PCR amplicons representative of different SSCP patterns revealed five different DNA sequences. All of these sequences showed a low sequence homology to known ovine *KRTAPs*, but were similar to *KRTAP13-n* sequences from humans and cattle, with the closest homologue being bovine *KRTAP13-3*. These ovine sequences were named *KRTAP13-3* alleles *A-E* and deposited in the GenBank with the accession numbers JN377429 - JN377433.

Among the sheep investigated, alleles *A, B* and *C* were common and occurred at frequencies of 28.9%, 35.0% and 32.0%, respectively. Alleles *D* and *E* were minor alleles and both were detected at a frequency of 2.0% in the sheep studied.

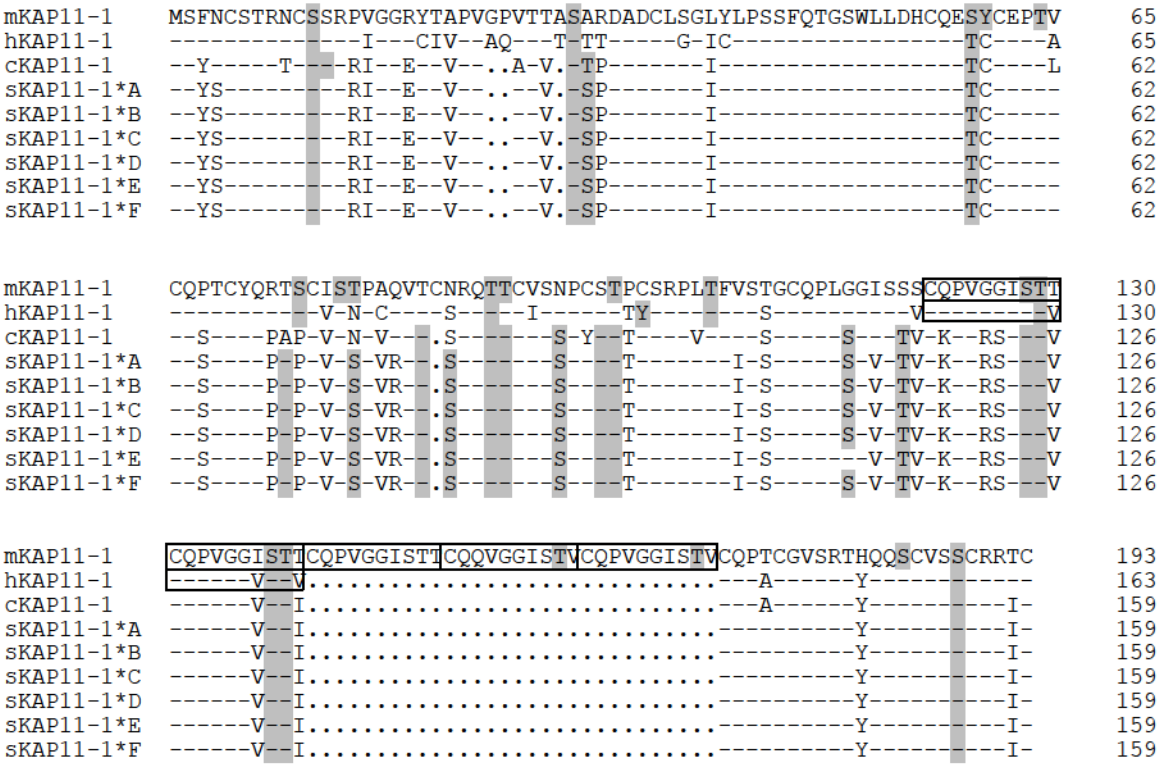


Figure 2-14. Alignment of the predicted amino acid sequences of the KAP11-1 genes. Amino acids identical to the top sequence are represented by dashes and dots have been introduced to improve the alignment. Decapeptide repeats are boxed, and potentially phosphorylated residues are shaded. The sheep sequences A-F (GenBank HQ595347 - HQ595352, respectively) are indicated with "s", while the sequences of human (GenBank NW_001838706), cattle (NM_001080740) and mouse (U03686) are indicated with "h", "c" and "m", respectively.

Four nucleotide substitutions were identified within the ovine *KRTAP13-3* coding region (Figure 2-15). Of these substitutions, three were non-synonymous and would result in amino acid changes (p.Arg79Cys, p.Arg81Gln and p.Tyr130His).

The ovine *KRTAP13-3* sequences all contained an open reading frame of 568 nucleotides, encoding a notional polypeptide of 156 amino acids. This polypeptide sequence was similar to both human (NM_181599.2, NM_181621.3, NM_181622.1 and NM_181600.1) and bovine KAP13-n sequences, but was more similar to bovine KAP13-3 (ENSBTAG00000040032) than bovine KAP13-1 (ENSBTAG00000003613) (Figure 2-16). The ovine KAP13-3 polypeptide contained a high level of serine (22.44 mol%), cysteine (11.54 - 12.18 mol%), arginine (10.26 - 11.54 mol%) and glycine (9.62 mol%), and these were the predominant amino acids, making up 53.82 - 55.78 mol% of the total amino acid content. Aspartic acid, glutamic acid, methionine and tryptophan were rare (accounting for 0.64 mol% each) and alanine was absent from the polypeptide.

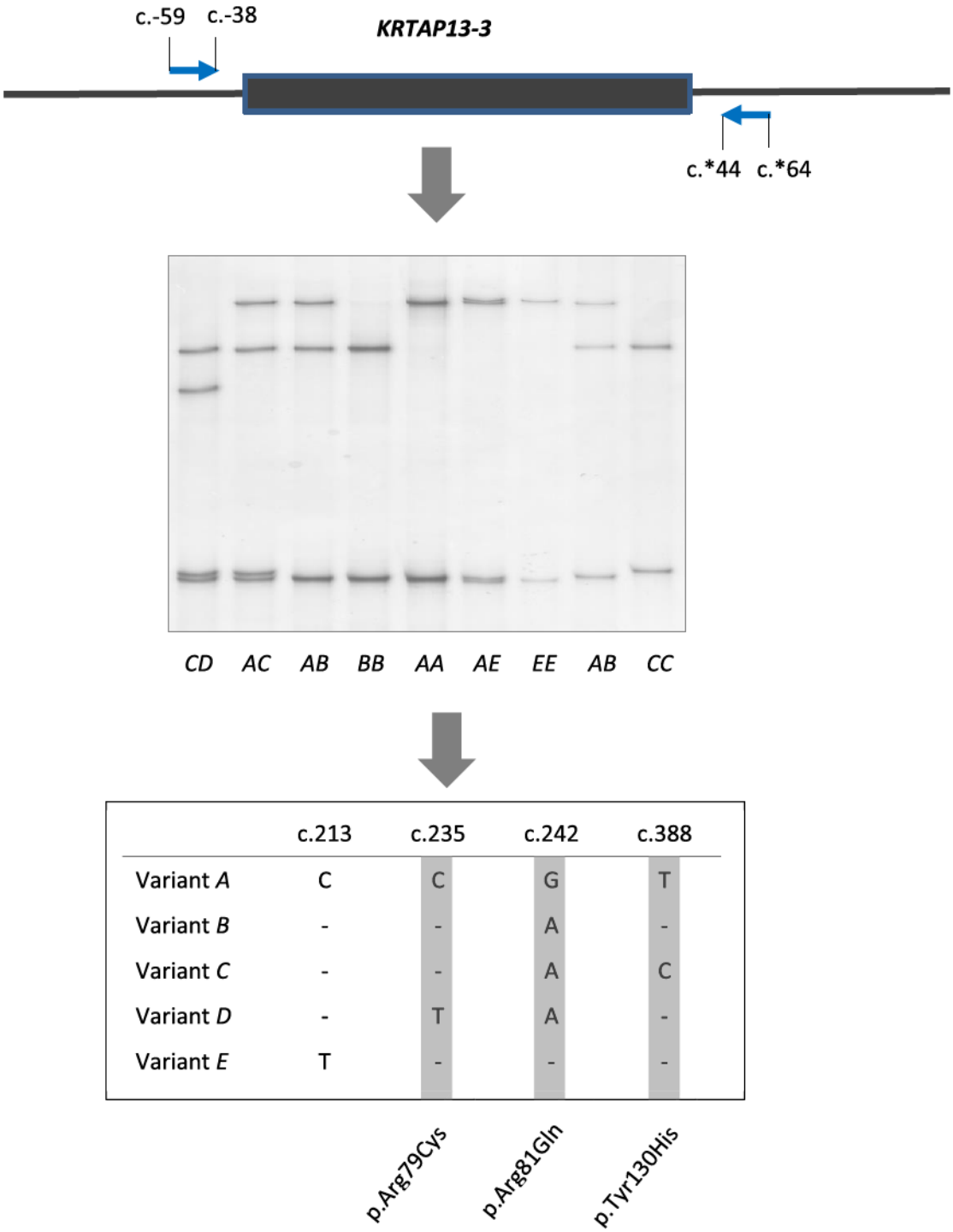


Figure 2-15. Variation identified in ovine *KRTAP13-3*. The entire coding sequence of *KRTAP13-3* was amplified using primers KRTAP13-3-up and KRTAP13-3-dn with the location of primers indicated. Five unique PCR-SSCP banding patterns representing five allelic variants (*A* to *E*) in either homozygous or heterozygous genotypes are shown. The non-synonymous SNPs identified in this study are shaded with the corresponding amino acid substitutions being shown. Nucleotide positions refer to GenBank JN377429 and the numbering of nucleotide and amino acids follow the HGVS nomenclature (den Dunnen & Antonarakis, 2000).

The notional polypeptide sequences had calculated pI values of 9.41, 9.26, 9.28, 9.04 and 9.41 for variants *A* to *E*, respectively. Between 18 to 21 serine and threonine resides in the notional ovine KAP13-3 polypeptide were predicted to be phosphorylated (Figure 2-16).

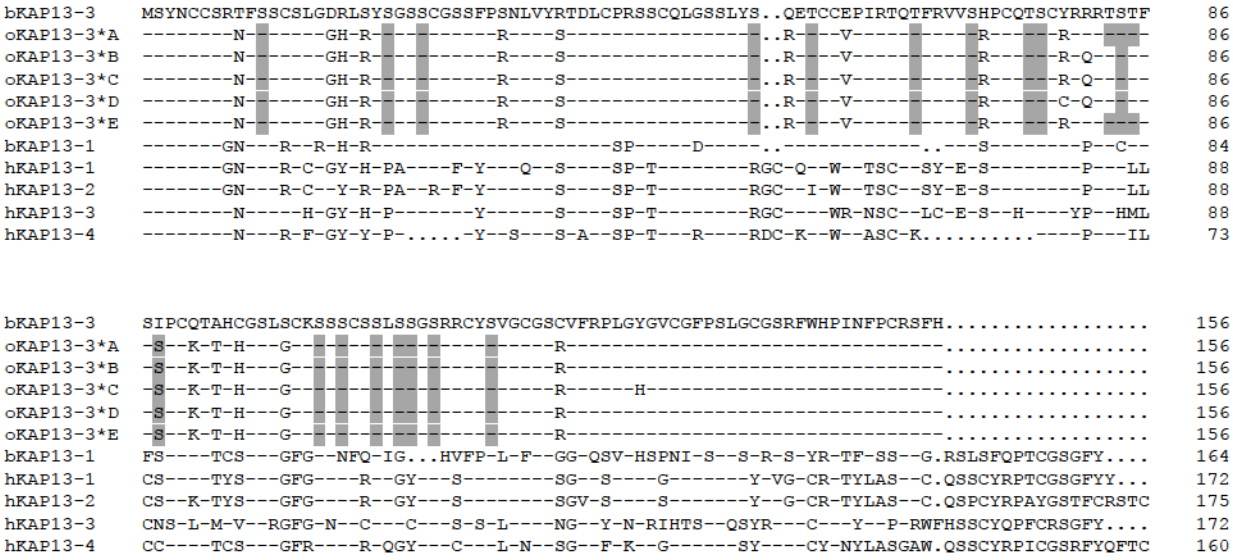


Figure 2-16. Alignment of the predicated amino acid sequences of the KAP13 gene from different species. Amino acids identical to the top sequence are presented by dashes, and dots have been introduced to improve the alignment. Potentially phosphorylated residues are shaded. The ovine sequences (JN377429 - JN377433 for alleles *A* - *E*, respectively) are indicated with a prefix "o", the bovine sequences (ENSBTAG00000003613 and ENSBTAG00000040032 for KAP13-1 and KAP13-3, respectively) are indicated with a prefix "b", and the human sequences (NM_181599.2, NM_181621.3, NM_181622.1 and NM_181600.1 for KAP13-1, KAP13-2, KAP13-3 and KAP13-4, respectively) are indicated with a prefix "h".

Identification of ovine *KRTAP24-1* with four allelic variants

This work has been published in "Zhou H, Gong H, Yan W, Luo Y and Hickford JG. Identification and sequence analysis of the keratin-associated protein 24-1 (KAP24-1) gene homologue in sheep. Gene 2012, 511:62-5".

A BLAST search of the Ovine Genome Assembly v2.0 using the human *KRTAP24-1*coding sequence (NM_001085455) revealed a homologous region on sheep chromosome 1 (OAR1:123817670_123818230; *E* = $8e^{-52}$). Analysis of the sequence in this homologous region led to the identification of a 759 bp open reading frame at OAR1:123817628_123818386. Six previously identified ovine KAP genes were also found near this open reading frame and these from closest to

farthest away were *KRTAP13-3*, *KRTAP6-1*, *KRTAP6-2*, *KRTAP8-1*, *KRTAP7-1* and *KRTAP11-1* (Figure 2-17).

Amplicons containing the complete open reading frame were produced using PCR and these were of the expected size (1003 bp). The amplicons exhibited variation upon SLCP analysis and four different SLCP patterns were detected (Figure 2-18). Either one PCR-SLCP pattern, or a combination of two patterns, was observed for each sheep.

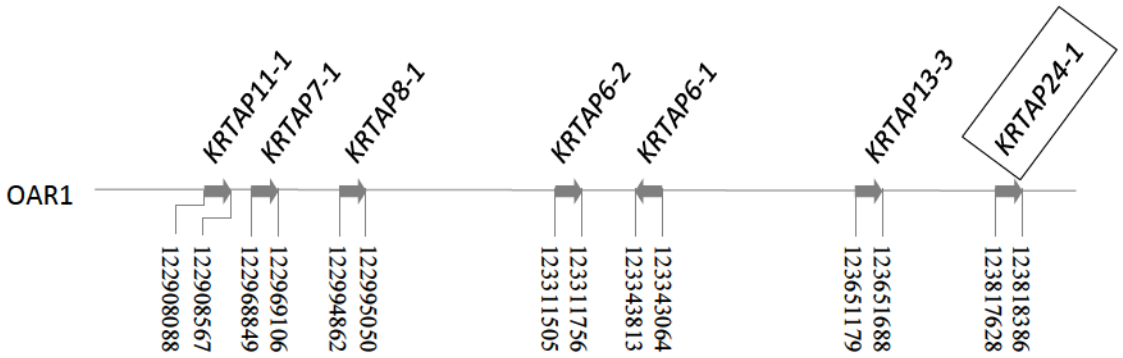


Figure 2-17. Location of seven KAP genes (*KRTAPs*) on sheep chromosome 1. The newly identified KAP24-1 gene is boxed. Arrow bars represent coding regions, with the direction of transcription being indicated. Nucleotide positions refer to the Sheep Genome Assembly v2.0.

Sequencing of amplicons representative of the four unique SLCP patterns revealed four different nucleotide sequences (Figure 2-18). All of these sequences were different to, but shared homology with, the sequence reported in the Ovine Genome Assembly v2.0. The differences likely reflecting either genetic variation or sequencing errors in the Assembly. These ovine sequences did not share high similarity with any other reported ovine *KRTAP* sequence. High sequence similarity was found with the human *KRTAP24-1* sequence (NM_001085455) and the predicted *KRTAP24-1* sequences from a number of species including cattle (XM_002684598.1), pig (XM_003358906.1), dog (XM_003433976.1), Sumatran orangutan (XM_002830616) and Northern white-cheeked gibbon (XM_003263833). These ovine sequences (named *A* to *D*) were concluded to represent allelic variants of *KRTAP24-1* in sheep, and were deposited into GenBank with accession numbers JX112014 - JX112017, respectively. Seven nucleotide substitutions were identified in the four sequences. All of these were located in the coding region and four of them were non-synonymous substitutions that would result in amino acid changes (Figure 2-18).

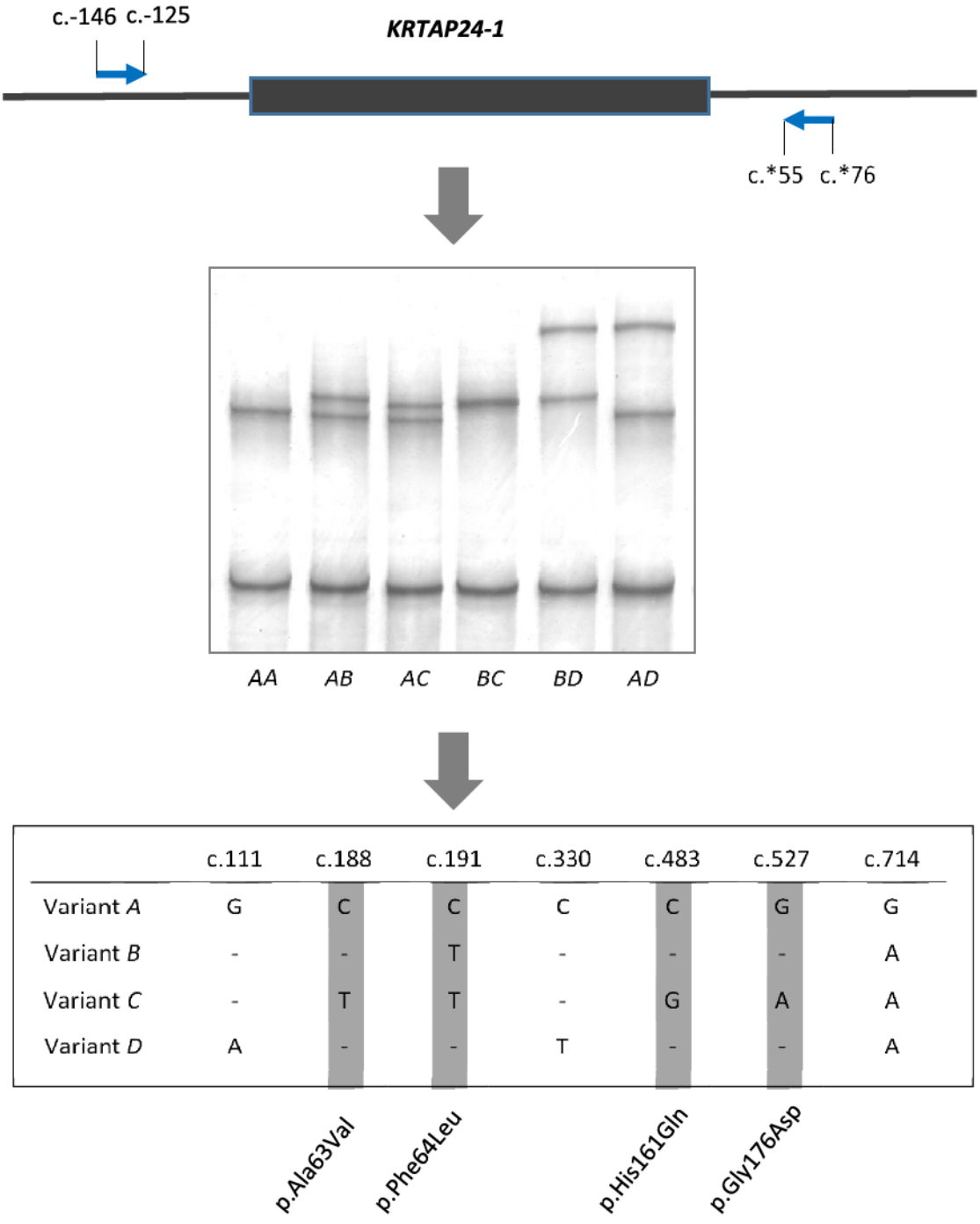


Figure 2-18. **Variation identified in ovine *KRTAP24-1.*** The entire coding sequence of *KRTAP24-1* was amplified using primers KRTAP24-1-up and KRTAP24-1-dn with the location of primers indicated. Four unique PCR-SLCP banding patterns representing four allelic variants (*A* to *D*) in either homozygous or heterozygous genotypes are shown. The non-synonymous SNPs identified in this study are shaded with the corresponding amino acid substitutions being shown. Nucleotide positions refer to GenBank JX112014 and the numbering of nucleotide and amino acids follow the HGVS nomenclature (den Dunnen & Antonarakis, 2000).

The putative ovine *KRTAP24-1* sequences encoded a polypeptide of 252 amino acid residues. This contained a high content of serine (14.7 mol%) and moderate levels of cysteine (8.3 mol%), glycine (8.3 mol%) and proline (8.0 - 8.3 mol%). Other residues that commonly occurred in the polypeptide included threonine, tyrosine, arginine and glutamine which accounted for 7.1, 6.8, 6.4 and 4.8 - 5.2 mol%, respectively. The calculated pI of the four putative polypeptides ranged from 8.4 to 8.5.

The putative polypeptides possessed two N-glycosylation sequons N-X-(S/T) (where X is any amino acid except proline) and seven potential O-glycosylation sites (Figure 2-19). In these polypeptides, either 21 (for variant *C*) or 22 (for variants *A*, *B* and *D*) residues were predicted to be potentially phosphorylated (Figure 2-19).

2.4 Discussion

2.4.1 Sequence and Length Variation in Previous Described KAP Genes

Ovine *KRTAP1-4* exhibits significant variation

This is the first study to report variation in ovine *KRTAP1-4*. The identification of nine allelic variants from 320 New Zealand sheep indicates that *KRTAP1-4* is highly polymorphic and as more sheep from more sire lines and breeds are investigated, it could be expected that more alleles may be found. The extent of the polymorphism found at this locus is higher than reported at any other *KRTAP* gene investigated with the exception of *KRTAP1-3* at which nine alleles have also been reported (Itenge-Mweza *et al.*, 2007; McLaren *et al.*, 1997; Rogers *et al.*, 1994b; Wood *et al.*, 1992). It is therefore conceivable that these nine alleles are in linkage with the nine alleles at the nearby *KRTAP1-3*, although that cannot be ascertained from these results.

Although the amount of variation found in *KRTAP1-3* and *KRTAP1-4* is comparable, the nature of variation appears to be different for these two genes. The *KRTAP1-4* SNPs tend to be clustered and most of the SNPs are non-synonymous. This is in contrast to *KRTAP1-3*, in which SNPs are spread across the entire gene and the majority of them (18 of 19) are synonymous (Itenge-Mweza *et al.*, 2007). This suggests that the polymorphism observed at these two loci may be derived from different mechanisms. The predominance of non-synonymous over synonymous SNPs at *KRTAP1-4* suggests that natural selection is, or has been acting on this gene (Nei, 2005).

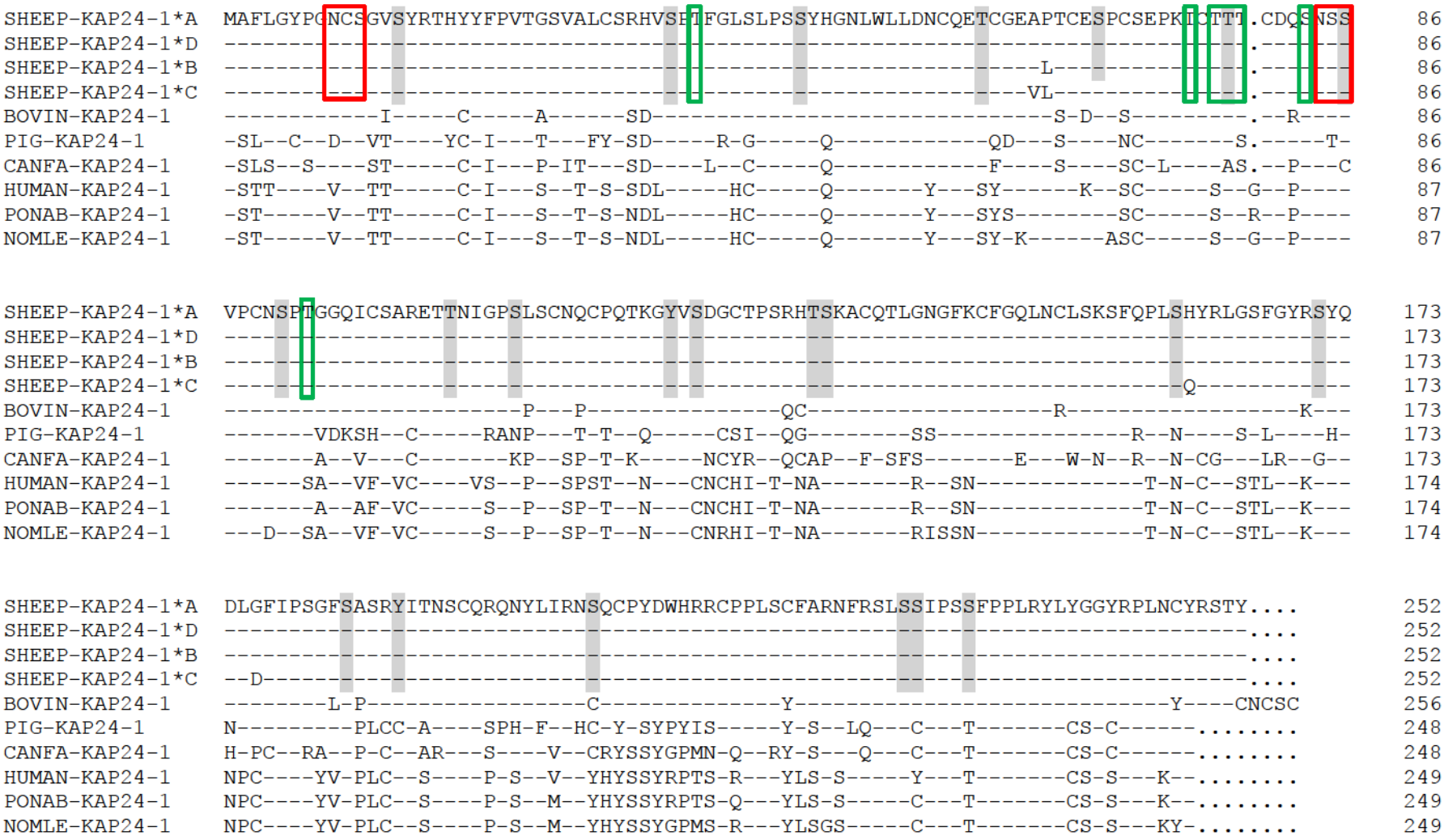


Figure 2-19. Alignment of the predicted KAP24-1 amino acid sequences. Amino acids identical to the top sequence are presented by dashes and dots are introduced to improve the alignment. In the sheep sequences, N-glycosylation sequons are indicated with red boxes, and potential O-glycosylation sites are shown in green boxes. The residues that are potentially phosphorylated are shaded.

It is observed that some *KRTAP1-4* alleles could be generated by exchanging short segments of DNA between two other alleles. For example, the exchange of a segment containing c.161 to c.273 between alleles *A* and *C* would generate allele *G*; allele *E* could be generated by exchanging a segment covering c.384 to c.*9 between alleles *A* and *I*; allele *H* could be generated by alleles *A* and *I* as a result of a segment exchange from c.236 to c.238. This suggests gene conversion has occurred in this gene and that it is one of the mechanisms generating the polymorphism.

Most (nine of 14) of the SNPs in ovine *KRTAP1-4* would lead to amino acid changes and may have an impact on the structure of the peptide encoded. In particularly, two SNPs at c.236 and c.491 would result in loss or gain of cysteine, which may affect cross-linking with the intermediate filaments or other KAPs. Substitutions at these positions may therefore affect wool characteristics.

Overall, ovine *KRTAP1-4* is a highly polymorphic gene and variation at this locus is likely to be in part responsible for the observed variation in wool characteristics. It could therefore be a candidate gene-marker for improving wool traits.

Polymorphism of the ovine *KRTAP5-4*

Variation in both sequence and length was observed in ovine *KRTAP5-4*. While length variation in the sequences encoding the cysteine-rich repeat segments has been described in KAP genes, it has only been reported previously in one family (KAP1) in sheep in which decapeptide length variation also occurs (Rogers *et al.*, 1994b) and two families (KAP1 and KAP4) in humans (Kariya *et al.*, 2005; Shimomura *et al.*, 2002). The detection of length variation in the ovine KAP5-4 gene suggests that variation in the number of repeats coding for cysteine-rich segments may be a structural hallmark of some of the KAP genes. This length variation would result in different levels of cysteine content at the protein level and given that cysteines in KAPs may be involved in the cross-linking with keratin IFs (Powell & Rogers, 1986), variation in the number of cysteine-rich repeats and variation in the amino acid sequence may affect the strength of the interaction, and hence wool or hair characteristics.

Diversity of ovine *KRTAP6*

The identification of more than two *KRTAP6* sequences from each sheep suggests that these sequences are derived from multiple loci and not from a single locus. This is supported by the results of sequence comparison which revealed that these sequences could be divided into three different groups based on sequence similarity.

The high sequence similarity of sequences *B* and *D* to a published ovine *KRTAP6-1* sequence suggests that these two sequences are allelic variants of ovine *KRTAP6-1*. The 57 bp deletion observed in Sequence *B* would result in a loss of 19 amino acid sequence, meaning sequence *B* would likely

encode the smallest known KAP protein. Length variation has been previously noted in the sheep KAP1 family (Rogers *et al.*, 1994b), and the human KAP1 and KAP4 families (Kariya *et al.*, 2005; Shimomura *et al.*, 2002). This suggests length variation may be also a structural hallmark of KAP6 genes.

The observation that sequences *A* and *C* were closely related to a partial sheep KAP6 amino acid sequence derived by protein sequencing (Gillespie, 1990), but different to other KAP6 sequences, suggests that sequences *A* and *C* are derived from the same locus that produces this partial KAP6 amino acid sequence which has been referred to as KAP6.2 (Fratini *et al.*, 1993).

While sequence *E* had low homology to both *KRTAP6-1* and nominal *KRTAP6-2* sequences, it exhibited higher homology to a putative *KRTAP6-3* sequence from cattle. This suggests that this sequence represent ovine *KRTAP6-3*. There was only one sequence identified (*E*), and this was detected in 11% of the sheep investigated. The absence of a *KAP6-3* amplicon in many individuals suggests either that some sheep lack the *KAP6-3* gene (referred to as a null allele), or it is possible that sequence variation may exist in the primer binding regions and hence some forms of the gene were not amplified under the PCR conditions used. Further investigation is required to confirm whether *KAP6-3* has a null allele, or whether un-identified new allele(s) exist, or a combination of these events. It suggests that *KAP6-3* is polymorphic in sheep.

Diallelic variation in *KRTAP7-1*

Only two alleles were detected in ovine *KRTAP7-1* and SSCP with different electrophoresis conditions did not reveal any further variation. The extent of variation detected here was somewhat lower than that reported by McLaren *et al.* (1997) using Southern hybridisation-RFLP. They reported two alleles using the restriction enzyme *Bgl*II and four alleles using *Msp*I in analysing three full-sib pedigrees of the AgResearch International Mapping Flock (IMF) (Crawford *et al.*, 1995). However, the sequence variation revealed here was not in a nominal *Bgl*II or *Msp*I recognition site. What is more, there are no *Bgl*II recognition sequences present or that can be created by a single nucleotide substitution, in the coding region of ovine *KRTAP7-1*. This suggests that these recognition sites were located outside of the coding region of the gene and hence the variation revealed by Southern hybridisation-RFLP was in regions flanking the gene. If this was the case, then one would expect only simple RFLP patterns when using a gene-specific probe, although this was not discussed by the authors. It is also possible that the discrepancy between our study and McLaren *et al.*'s (1997) may be due to different samples being investigated. Nevertheless, this study shows that PCR-SSCP detected sequence variation in the ovine *KRTAP7-1* coding region that cannot be detected by Southern hybridisation-RFLP in the earlier study, as the latter is incapable of detecting sequence substitutions located outside of restriction endonuclease recognition sites.

Variation in *KRTAP8-1* is lower than expected

The variation reported here in the *KRTAP8-1* coding region appears to be less than that reported by Wood *et al.* (1992) who reported eleven *KRTAP8-1* alleles in 33 unrelated sheep from various breeds (including the NZ Romney, Coopworth, and Perendale breeds, and various hybrids of these breeds) based on dinucleotide length variation in PCR amplicons of the 5' flanking region of the gene. In that study, the PCR primers were designed based on the published ovine *KRTAP8-1* sequence (GenBank X05639) and were expected to amplify an amplicon of 124 bp. However, the PCR amplicons reported by Wood *et al*. (1992) were 152 - 175 bp long, and did not match the expected size (124 bp). There was no sequencing data reported for the amplicons by these authors and accordingly the sequence of the nominal 5' flanking region variation awaits validation.

While most of the SNPs identified in ovine *KRTAP8-1* were silent; silent nucleotide substitutions may affect mRNA stability (Duan *et al.*, 2003) and accordingly all the variation detected here may affect the structure and/or expression of KAP8-1, as may that variation reported by Wood *et al.* (1992). In both cases this genetic variation may affect wool traits.

2.4.2 Newly Identified KAP Genes

Ovine *KRTAP1-2* is a newly identified and polymorphic gene within the KAP1 family

This is the first study that provides evidence for the presence of a gene encoding KAP1-2 (Swiss-Prot P02439.1; B2B or SCMK-B2B) on sheep chromosome 11 where the other KAP1-n genes have been located previously (McLaren *et al.*, 1997). This gene possesses a small, intronless reading frame and is located between the KAP1-1 and KAP1-3 genes which are approximately 11 kb apart. This is consistent with the notion that the KAP genes are typically small and clustered together.

In its coding region, the ovine KAP1-2 gene shares a high level of sequence homology with the other known members of the ovine KAP1-n family, and the notional encoded polypeptide possesses a number of sequence motifs that are unique to, and conserved within, the KAP1-n family. This provides good evidence that the gene identified is a member of this family. Additionally, unique DNA sequences found in both the 5' and 3' flanking regions strongly support the contention that there are at least four different and unique genes within the KAP1 family and that the *KRTAP1-2* sequences obtained represent a newly identified KAP1-n gene and not allelic variants of *KRTAP1-1*, *KRTAP1-3* or *KRTAP1-4*. This was confirmed by the mapping of all four KAP1-n genes onto a defined region of sheep chromosome 11 (Figure 2-7).

While KAP1-2 is very similar to the other known family members (KAP1-1, KAP1-3 and KAP1-4) of the KAP1-n family, it differs in a number of respects (Figure 2-8). Firstly, it has unique sequence motifs.

Each member of KAP1 appears to possess some unique sequences, and a unique pattern of sequence motifs can be identified for each member. Secondly, individual KAP1 members can be partially differentiated by the number of decapeptide "QTSCCQPXXX" repeats in the central repetitive region. There are between three and five repeats in KAP1-1 (Itenge-Mweza *et al.*, 2007; Rogers *et al.*, 1994b), whereas for other members the repeat numbers are not variable. KAP1-2 has three repeats, compared to two repeats in KAP1-3 (Itenge-Mweza *et al.*, 2007) and five repeats in KAP1-4 (Gong *et al.*, 2010b). Thirdly, there is an apparent loss of a five residue C-terminus tail (encoded by 15 nucleotides) in KAP1-2, compared to other KAP1 proteins. Although the KAP1-n genes exhibit a high level of sequence similarity in the coding region, they are quite different outside this region and sequence similarity appears to be only observed in three small regions. One conserved region contains the likely TATA box and the other two conserved regions are located immediately 5' and 3' to the coding region.

The *KRTAP1-2* alleles found in the sheep genomes confirmed the amino acid variation (p.Gly78Asp) reported previously (Elleman & Dopheide, 1972). Additionally other variation in the gene appears to be spread throughout the gene, and the majority of nucleotide substitutions (6 out of 8) in the coding region were synonymous. This trend is similar to that reported for *KRTAP1-3* (Itenge-Mweza *et al.*, 2007), but different to that found in *KRTAP1-4* in which the nucleotide substitutions appear to be predominantly clustered and non-synonymous (Gong *et al.*, 2010a). This suggests the genetic variation in *KRTAP1-2* and *KRTAP1-3* may originate by a similar mechanism, but that variation accumulates for a different reason in *KRTAP1-4*. This is also consistent with the observation that *KRTAP1-2* is physically closer to *KRTAP1-3*, than *KRTAP1-4* on chromosome 11 (Figure 2-7).

It is interesting to note that to date, *KRTAP1-2*, *KRTAP1-3* and *KRTAP1-4* appear to possess the same number of alleles, with nine described for each (Gong *et al.*, 2010a; Itenge-Mweza *et al.*, 2007; Rogers *et al.*, 1994b). As the KAP1-n genes are clustered on sheep chromosome 11 (Figure 2-7), it is conceivable that the alleles found in these genes may be in linked into an extended haplotype. Considering that *KRTAP1-1* is also clustered with these genes and located between *KRTAP1-4* and *KRTAP1-2*, it is possible that *KRTAP1-1* may exhibit a comparable level of variation, although only three alleles have been reported. To date only length variation has been described in *KRTAP1-1* (Itenge-Mweza *et al.*, 2007; Rogers *et al.*, 1994b), although sequence variation has not been looked for.

The identification of *KRTAP1-2* brings the number of KAP1-n genes in the sheep genome from three to four. While the number of KAP1-n genes found in sheep and human are comparable, there are some differences in individual KAP1-n genes between these two species. Firstly, all the KAP1-n genes are polymorphic in sheep (Gong *et al.*, 2010b; Gong *et al.*, 2011e; Itenge *et al.*, 2010; Rogers *et al.*,

1994b), but in the human only two of the genes are currently known to be variable (Shimomura *et al.*, 2002). Secondly, the level of polymorphism found in sheep is much higher than that reported in humans. In sheep, up to nine alleles have been found in all of the KAP1-n genes except the KAP1-1 gene, when to date there are maximally three allelic variants reported for the human KAP1-n genes (Shimomura *et al.*, 2002). Thirdly, gene length variation appears to be more common in humans than in sheep. Of the two human KAP1-n genes that are multiallelic, both exhibit length variation (Shimomura *et al.*, 2002), but in sheep only *KRTAP1-1* has been found to vary in length (Rogers *et al.*, 1994b). These differences between the human and sheep genes suggest that the direct mapping of KAP genes across species may be challenging and any future effort to define a unifying nomenclature for the KAPs will need to be cognisant of this.

The sheep KAP 8-2 gene, a new KAP8 family member that is absent in humans

This study has identified a new gene encoding a HGT-KAP in sheep. The gene was grouped with other KAP genes on ovine chromosome 1, but located at a different position and with a lower sequence similarity to these genes. This suggests that this gene represent a previously unidentified ovine KAP gene. The similarity of this gene sequence to the *KRTAP8-2* sequences from goats and reindeer suggests that it is an ovine orthologue of *KRTAP8-2*.

The putative ovine *KRTAP8-2* exhibited sequence variation, with two sequence variants being found. This is consistent with the finding of sequence variation in other ovine *KRTAPs* (Gong *et al.*, 2011b; Gong *et al.*, 2011c; Gong *et al.*, 2012; Zhou *et al.*, 2012). However, in contrast to other *KRTAPs*, the variation found in ovine *KRTAP8-2* was not within the coding region, but instead located near the TATA box. This variation may affect RNA polymerase II binding and hence the expression of the gene, but this would need to be confirmed through further functional investigations.

The predicted ovine KAP8-2 amino acid sequence exhibits some characteristics that are consistent with other type II HGT-KAPs, such as a high glycine and tyrosine content and higher levels of phenylalanine (Table 2-3). However, some unique features are also observed. Firstly, there is a relatively low cysteine content (3.2 mol%), which contrasts with all previously reported KAPs. Secondly the polypeptide contains a high (4.8 mol%) aspartic acid and glutamic acid content. These acidic amino acids are not common in other HGT-KAPs. Lastly it is noteworthy that the polypeptide would likely have a low pI (6.3), as a result of this relatively high level of acidic amino acid residues. Such a low pI value has not been observed in any other HGT-KAP, where the pI is typically higher than 8.

Considering there are two types of keratins that cross-link with the KAPs, and of these the type I keratins are characteristically more acidic (pI 4.5 - 6.0), while the type II keratins tend to be more

basic (pI 6.5 - 8.5) (Bowden *et al.*, 1987). The predicted lower pI value of KAP8-2 may affect its interaction with keratins, and on a charge basis it would be expected to have a greater affinity for the type II (basic) keratins.

While the protein encoded by the ovine KAP8-2 gene has not yet been isolated from wool, the gene appears to be expressed and functional in sheep as many ESTs with sequences identical to this gene have been reported in skin tissues (Table 2-2). A functional orthologue of this gene appears to be absent in humans, a species in which only one functional and two pseudogenic KAP genes are found (Rogers *et al.*, 2002). The KAP8-2 gene is the only KAP gene identified and reported to date that is present in sheep and goats, but is absent in humans. The functional significance of this gene in hair and wool characteristics, and in the evolution of hair and wool, awaits further investigation.

Identification of ovine *KRTAP11-1*

This is the first study describing *KRTAP11-1* in sheep. The ovine sequences do not show high homology to any other known ovine KAP gene sequence, suggesting that they are derived from a previously unidentified gene and are not allelic variants of known KAP genes. This is supported by the observation that all the sheep investigated contained either one or two sequences, which is consistent with them being either homozygous or heterozygous at this locus. The closest homologies with these ovine sequences are found in the KAP11-1 gene sequences of other species, which suggests that they are the ovine ortholog of the KAP11-1 gene.

While KAP11-1 sequences from different species share a moderately high sequence similarity, there are some obvious differences between species. First is the occurrence of a decapeptide repeat structure. Mouse KAP11-1 exhibits a strong decapeptide repeat structure (Huh *et al.*, 1994), but this repeat structure appears to be weak in humans (Rogers *et al.*, 2002) and absent in both sheep and cattle (Figure 2-14). The second difference is in the length of the polypeptide. Mouse KAP11-1 is the longest ortholog containing 193 amino acids, followed by the human ortholog with 163 amino acids. KAP11-1 from sheep and cattle are shortest, each containing 159 amino acids.

The ovine KAP11-1 gene, like its orthologs from human (Rogers *et al.*, 2002) and mouse (Huh *et al.*, 1994), contains a single exon or intron-less reading frame, encoding a cysteine-rich protein. The abundance of cysteine, serine and threonine, together with low concentrations of histidine, lysine and methionine in the KAP11-1 protein is consistent with the assignment of KAP11-1 into HS-KAPs (Rogers *et al.*, 2002). KAP11-1 is however unique in two respects. Firstly, the protein contains a low frequency of the "CCXP" motif, which is common in the cysteine-rich KAPs and is often found as part of longer repeats (Powell & Rogers, 1997). This motif is present as a single copy in KAP11-1 from sheep, cattle and human, and is absent in mouse KAP11-1. Secondly, KAP11-1 does not have a high

content of glycine, but instead has more serine and threonine residues. These two amino acids make up over 30 mol% of the residues in ovine KAP11-1. The biological function of having high concentrations of serine and threonine in KAP11-1 is unknown. However, as many of the serine and threonine residues are predicted to be potential phosphorylation sites, and the number of potential phosphorylation residues is two to four times higher than any other HS-KAP, it may be associated with the phosphorylation of this KAP. While protein phosphorylation has not yet been described in KAPs, it has been reported in keratins and is known to affect keratin assembly state and organization (Ku *et al.*, 1996). Despite the high frequency of serine and threonine residues, these do not occur in the glycosylation sequence motif N-X-T/S, especially as asparagine is uncommon in the putative protein. While the potential might exist for O-linked glycosylation, there is no evidence of KAPs being glycosylated in the keratinized wool fibre.

The identification of six alleles of ovine *KRTAP11-1* is consistent with the trend of many ovine *KRTAPs* being polymorphic (Gong *et al.*, 2010a; Itenge-Mweza *et al.*, 2007; Rogers *et al.*, 1994b). While the majority of sequence variations identified in ovine *KRTAP11-1* were silent, it is interesting to note that some variation may result in the gain or loss of potential phosphorylation sites. Equally, the functional effects of silent variation should not be ignored because they may either act alone or together to influence mRNA stability and translation (Duan *et al.*, 2003), or they may affect expression through changing codon-preference.

Of the five nucleotide substitutions identified in ovine *KRTAP11-1*, three (including the non-synonymous substitution) were found in *KRTAP11-1* sequences from other species, suggesting a discrete range of variation is tolerated, or alternatively that variation in these genes arose prior to species divergence. It would also suggest the substitutions are unlikely to be the result of sequencing or amplification errors.

Overall the genetic variation found in *KRTAP11-1* may influence its expression, protein structure, and/or post-translational modifications, and consequently affect wool fibre structure and wool traits.

Identification of ovine *KRTAP13-3*

This study has identified a notional KAP13-3 gene in sheep and investigated the variation in its coding region. The newly identified gene was intronless and contained an open reading frame encoding a cysteine-rich (11.54 - 12.18 mol%) polypeptide. These characteristics are consistent with other HS-KAPs.

Some unique features were observed within the notional ovine KAP13-3 gene sequence. Firstly, the putative polypeptide sequence had a high content (over 28 mol%) of serine and threonine

collectively, with many of these residues (41 - 45%) notionally being able to be phosphorylated. This pattern is similar to ovine *KRTAP11-1* (Gong *et al.*, 2011b), but different from other known HS-KAP genes. While a comparable level of serine and threonine is also found in the KAP1 family members, with the exception of KAP1-4, the proportion of these two amino acids that could potentially be phosphorylated in the putative KAP13-3 is two to three times higher than those KAP1 family members. While phosphorylation has not yet been investigated in KAPs, Ku *et al.* (1996) reported that phosphorylation in keratins can affect keratin assembly and organization.

The putative ovine KAP13-3 appeared to be a basic protein. It contained many more (14.3 - 15.5 mol%) positively charged amino acids (arginine, lysine and histidine) than negatively charged ones (aspartic acid and glutamic acid; 1.2 mol% in total), resulting in a high (9.0 - 9.3) calculated pI value. Such a high pI value has not been observed for any other known HS KAP, where the pI values are typically less than 7. There are two types of keratins that cross-link with KAPs, and of these, the type I keratins are typically more acidic (pI 4.5 - 6.0), while the type II keratins tend to be more basic (pI 6.5 - 8.5) (Bowden *et al.*, 1987). The high pI value of KAP13-3 may accordingly affect its interaction with keratins, and on a charge basis alone it would appear to have a greater affinity for the type I (acidic) keratins.

The ovine sequences identified were more homologous to bovine *KRTAP13-3* than to bovine *KRTAP13-1* (the only other known family member in cattle), suggesting that these sequences represent the ovine *KRTAP13-3* locus. However, while ovine *KRTAP13-3* showed a reasonable degree of homology to human *KRTAP13-n* sequences, it did not cluster to any particular human *KRTAP13-n* gene. This suggests that genes of the KAP13 family emerged before the divergence of sheep and cattle, but after the divergence of primates and the artiodactyl mammals. Investigation of this gene and its homologues from other species could shed further light on the evolution of *KRTAPs*.

Four nucleotide substitutions were identified in the ovine *KRTAP13-3* coding region, and three of them would result in amino acid changes. It is interesting to note that some amino acid changes may result in the gain or loss of potential phosphorylation sites and also result in variation in the net charge of the protein. This variation in ovine *KRTAP13-3* may affect the expression, structure and assembly of the protein and consequently influence key wool traits that underpin wool's quality and value.

Identification of ovine *KRTAP24-1*

The putative ovine KAP24-1 gene was identified on chromosome 1. The gene was clustered with several previously described KAP genes and displayed a lower sequence similarity to any known ovine KAP gene when compared to the KAP24-1 gene from other species including humans, cattle,

dog, pig, Sumatran orangutan and Northern white-cheeked gibbon. This suggests that the gene represent the ovine *KRTAP24-1* sequence.

While the polypeptide encoded by the ovine KAP24-1 gene was rich in cysteine and could be classed into the high-sulphur KAP group, the cysteine content (8.3 mol%) was lower than in any other high-sulphur ovine KAP protein. In contrast, the ovine KAP24-1 protein contained a relatively high content of serine (14.7 mol%). Having a higher content of serine residues than cysteine residues is not common in high and ultra-high sulphur KAPs, but it has been described previously in the high sulphur proteins KAP11-1 (Gong *et al.*, 2011b) and KAP13-3 (Gong *et al.*, 2011c) for which the encoding genes have also been located near the gene on chromosome 1. Serine residues can be phosphorylated, so having large amounts of serine may suggest that KAP24-1 is phosphorylated. While phosphorylation has not yet been detected in KAPs, it has been reported in keratins and it has been shown to affect their assembly and organisation (Ku *et al.*, 1996).

The putative ovine KAP24-1 polypeptide contained a high level of tyrosine (6.8 mol%), and this is not seen in other high or ultra-high sulphur KAPs. This high level of tyrosine was not part of dimeric glycine-tyrosine repeats, which are characteristically seen in high glycine-tyrosine KAPs (e.g. the KAP6 family).

It is interesting to note that the ovine KAP24-1 protein may be glycosylated, as it possessed potential N-glycosylation and O-glycosylation sites. However, the polypeptide does not appear to have a signal peptide and proteins without signal peptides are unlikely to be exposed to the glycosylation mechanisms located in the endoplasmic reticulum and the Golgi apparatus. There is however emerging evidence that O-glycosylation also occurs within the nucleus and cytoplasm (Hart & West, 2009) and therefore the possibility of glycosylation of this polypeptide cannot be ruled out. This will require further investigation.

O-glycosylation has been detected previously on keratins K8 and K18 (Srikanth *et al.*, 2010) and it appears to affect keratin solubility and assembly in a way that is similar to that reported for phosphorylation (Srikanth *et al.*, 2010). Glycosylation and phosphorylation may occur on the same, or at proximal residues in ovine KAP24-1, suggesting there may be a relationship between these post-translational modifications.

The ability to detect variation in a 1003 bp fragment of *KRTAP24-1* supports our previous finding that PCR-SLCP is capable of detecting sequence variation in larger DNA fragments of up to 1 kb in size (Zhou *et al.*, 2011). By comparison PCR-SSCP analysis is better suited to small fragments in the range of 150 - 600 base pairs (Gong *et al.*, 2011e; Nataraj *et al.*, 1999; Zhou *et al.*, 2007).

The sequence variation identified is consistent with the trend of *KRTAPs* being polymorphic (Gong *et al.*, 2011a; Gong *et al.*, 2011c; Gong *et al.*, 2011e; Gong *et al.*, 2012; Nataraj *et al.*, 1999; Zhou *et al.*, 2007). Four of the nucleotide substitutions would result in amino acid changes and three of these amino acid changes occurred at positions that are conserved across species (Figure 2-19). Some of these amino acid substitutions may result in the gain or loss of potential phosphorylation sites. The variation described in ovine *KRTAP24-1* may affect the structure and assembly of the protein and consequently influence wool traits.

The research in sheep presented in this chapter and the research in humans indicate extensive diversity of KAPs, not only in terms of gene number but also in terms of genetic variation. To better accommodate this emerging diversity in KAP genes, the current KAP nomenclature was revised and an updated nomenclature is proposed in Chapter 3.

The results in this chapter also suggest that the extent of KAP diversity in sheep is probably higher than that found in humans. All the genes that are investigated and present in humans can be found in sheep, including *KRTAP1-2* (was historically named as *KRTAP1-3* in humans), *KRTAP11-1*, *KRTAP13-3* and *KRTAP24-1*. Sheep also carry an additional gene (*KRTAP8-2*) that appears to be absent in humans. Furthermore, all the KAP genes investigated are polymorphic in sheep, and the level of polymorphism is also higher than that reported in humans. This indicates that the significance of variation in ovine *KRTAPs* warrants further investigation. The effect of variation in some *KRTAPs* on wool traits was therefore investigated and this is presented in Chapter 4 of this thesis.

Chapter 3

An Updated Nomenclature for the Keratin-associated Proteins (KAPs) and the Genes that Encode them

This work has been published in "Gong H, Zhou H, McKenzie GW, Hickford JGH, Yu Z, Clerens S, Dyer JM and Plowman JE. (2010). Emerging issues with the current keratin-associated protein nomenclature. International Journal of Trichology 2:104-5" and "Gong H, Zhou H, McKenzie GW, Yu Z, Clerens S, Dyer JM, Plowman JE, Wright MW, Arora R, Bawden CS, Chen Y, Li J and Hickford JGH. (2012). An updated nomenclature for keratin-associated proteins (KAPs). International Journal of Biological Sciences 8:258-64".

3.1 Introduction

The research presented in the previous chapter indicates extensive diversity of KAPs, not only in terms of gene number but also in terms of genetic variation. To better accommodate this emerging diversity in KAP genes, the current KAP nomenclature was revised and an updated nomenclature was proposed.

The nomenclature system for KAPs has not been reviewed since the proposition and adoption of the system of Rogers and Powell. This was first proposed in 1993 (Rogers & Powell, 1993) and explained in detail in 1997 (Powell & Rogers, 1997). While this system has served us well until now; the differences now being described between species, the extensive genetic variation now being documented in some *KRTAPs* and some misuse or misinterpretation of the nomenclature, suggested the need for revision.

3.2 Past and Current Nomenclatures for Keratin-associated Proteins (KAPs)

The earliest attempts to classify keratins had their origins in the methods used to separate wool proteins. In 1934 these proteins were divided into two extractable classes: those with a lower sulphur content than whole wool, and those with a higher sulphur content, otherwise known as SCMK-A and SCMK-B respectively (Goddard & Michaelis, 1935). This division was based on the fractional 'salting-out' of s-carboxymethylated proteins. Subsequently the former group became known as the intermediate filament proteins, while the latter group became the KAPs. The advent of amino acid analysis enabled a further sub-division of the high sulphur class (SCMK-B) into high (HS) and ultra-high sulphur (UHS) proteins, with this split being based on whether their cysteine content

was above or below 30 mol% (Gillespie, 1966; Gillespie & Broad, 1972). Amino acid analysis also led to the discovery of a third class of proteins in wool that proved to be rich in glycine and tyrosine, the so-called high glycine-tyrosine proteins (HGT) (Harrap, 1963; Stein & Moore, 1948).

Subsequent attempts to fractionate the HS group of proteins and identify sub-components led to further improvement in our understanding of this class of proteins and also a proliferation of new protein names. The use of fractional precipitation with ammonium sulphate solutions resulted in the definition of two fractions SCMK-B1 and SCMK-B2 (Gillespie, 1963), with subsequent sub-fractionation of the B2 group into a further three components (their names now shortened to B2A, B2B and B2C) by chromatography (Lindley & Elleman, 1972). In parallel with these studies, column electrophoresis was used to fraction the HS components; one component, SCMK-BIII, being split by gel filtration into two new HS protein families: BIIIA and BIIIB (Haylett *et al.*, 1971; Swart *et al.*, 1969).

The HGT were sub-divided into the Type I and II sub-classes by ion-exchange chromatography (Gillespie & Darskus, 1971), the former being of moderate percentage of glycine and tyrosine, and comprising two components C2 and F (Gillespie, 1983). In contrast, the Type II family proteins, which contain a higher percentage of these two amino acids, were thought to contain up to 10 individual components (Powell & Rogers, 1994; Powell & Rogers, 1997), although only one has been fully sequenced to date (Fratini *et al.*, 1993). Finally, although their existence had been known about for some time, members of proteins from the UHS group were identified, the first cuticle UHS proteins being sequenced in 1990 and 1994 (Jenkins & Powell, 1994; MacKinnon *et al.*, 1990) and cortical UHSs in 1994 and 1995 (Fratini *et al.*, 1994; Powell *et al.*, 1995).

The increasing diversity of the KAPs, coupled with their non-uniform naming, led Rogers and Powell (Powell & Rogers, 1997) to suggest a nomenclature for the KAP proteins and genes using the abbreviation KAPm.nxpL for the protein and *KRTAPm.nxpL* for the gene. In subsequent iterations of this system the gene name became italicised, although use of this convention is only sporadic in the literature. In Rogers and Powell's system (Powell & Rogers, 1997), "m" denotes a family or unique protein, "n" denotes a component number, "x" denotes a variant, "p" denotes a pseudogene and "L" stands for "like". This nomenclature divides the KAPs of all species into families and further into family members based on similarities in their amino acid sequences. Historically then, what was originally called SCMK-B became SCMK-B2, then HS-B2A and then KAP1.1 for the protein and *KRTAP1.1* for the gene. Somewhat strangely through this time, the HS, UHS and HGT classification system persisted, perhaps reflecting that the abbreviations gave some indication of the type of protein being described, although in the last few years the discovery of KAPs that contain moderate amounts of cysteine and glycine has made this older classification system even more inadequate.

3.3 Problems Encountered with the Current Nomenclature

A variety of issues have arisen since 1997, which lead us to propose that the KAP/*KRTAP* nomenclature should be revised and adjusted. Firstly, the 2006 release of a consensus nomenclature for the mammalian keratins (Schweizer *et al.*, 2006) identified the need to adhere to guidelines proposed by the Human Genome Nomenclature Committee (HGNC), whose prominence in the area of nomenclature grew following the sequencing of the human genome. Consistency with the recommendations of this organisation seemed sensible, especially in the context of what was known about KAPs/*KRTAPs*. While the Powell and Rogers' (1997) system was useful, the protein nomenclature included a term "p" for a pseudogene, a term that could only really describe a non-expressed or faulty form of a gene and not a protein. The gene nomenclature was also somewhat confusing as "p" and "L" probably should not be present together. There was also some confusion over the use of punctuation in gene names, with both full stops and hyphens being used between the m (family) and n (constituent) in the nomenclature and with seemingly little consistency or pattern.

HGNC suggests punctuation should be avoided, the exception being its use in defining groups of related genes and in this respect, in this proposed nomenclature it was attempted to define a system that would lead to greater consistency in naming.

Two other and more substantive issues had also emerged, the increasing diversity of the genes from different species and the genetic variation there-in. It was felt that these matters needed to be better accommodated in the nomenclature.

To date, more than 100 KAP genes have been isolated from a range of mammalian species including sheep, humans, mice and rabbits. These genes have been placed in 27 families, each comprising 1 - 12 members (Rogers & Schweizer, 2005; Rogers *et al.*, 2008; Rogers *et al.*, 2006; Rogers *et al.*, 2007). In the human genome, the 89 functional KAP genes identified have been placed into 25 families (Table 3-1), although Wu *et al.* (2008) suggest up to 122 functional (n = 101) and pseudo (n = 21) genes based on analysis of sequences lodged in databases. These genes are clustered into five domains on three different human chromosomes (Table 3-2). In an analysis of eight species, Wu *et al.* (2008) have suggested that humans have a similar number of genes to other primates, but that rodents have an expanded repertoire. While there appears to be conservation at the sequence level across species (Wu *et al.*, 2008), it is still not clear as to which of these genes are expressed and where and when this expression occurs. In humans, all of the KAP genes are reported to be expressed in the hair follicle except for families 16, 22, 25 and 27 (Rogers & Schweizer, 2005; Rogers *et al.*, 2007).

In sheep, only 23 functional *KRTAPs* from 11 families have been reported to date and homologues for the other human genes have not been identified yet (Table 3-1). This is probably, in part, a result of the limited amount of research undertaken on sheep *KRTAPs* in the last decade.

Table 3-1. KAP genes identified in humans and sheep

	KAP family	Members identified in humans[†]		Members identified in sheep[†φ]	
HS	KAP1	4+2	(Rogers *et al.*, 2001)	4	(Gong *et al.*, 2011e; Powell *et al.*, 1983)
	KAP2	5+1	(Rogers *et al.*, 2001)	1	U60024, unpublished
	KAP3	3+1	(Rogers *et al.*, 2001)	3+1	(Frenkel *et al.*, 1989)
	KAP10	12+1	(Rogers *et al.*, 2004b)		
	KAP11	1	(Rogers *et al.*, 2004b)	1	(Gong *et al.*, 2011b)
	KAP12	4+1	(Rogers *et al.*, 2004b)		
	KAP13	4+2	(Rogers *et al.*, 2004b)	1	(Gong *et al.*, 2011c)
	KAP15	1	(Rogers *et al.*, 2004b)		
	KAP16	1	(Rogers *et al.*, 2001)		
	KAP23	1	(Rogers *et al.*, 2007)		
	KAP24	1	(Rogers *et al.*, 2007)	1	(Zhou *et al.*, 2012)
	KAP25	1	(Rogers *et al.*, 2007)		
	KAP26	1	(Rogers *et al.*, 2007)		
	KAP27	1	(Rogers *et al.*, 2007)		
UHS	KAP4	11+1	(Rogers *et al.*, 2001)	2	(Fratini *et al.*, 1994; Yu *et al.*, 2009)
	KAP5	12+2	(Yahagi *et al.*, 2004)	4+1	(Jenkins & Powell, 1994; MacKinnon *et al.*, 1990; Powell & Rogers, 1986)
	KAP9	7+1	(Rogers *et al.*, 2001)		
	KAP17	1	(Rogers *et al.*, 2001)		
HGT	KAP6	3	(Rogers *et al.*, 2002)	3	(Fratini *et al.*, 1993; Gong *et al.*, 2011a)
	KAP7	1	(Rogers *et al.*, 2002)	1	(Kuczek & Rogers, 1987)
	KAP8	1+2	(Rogers *et al.*, 2002)	2	(Gong *et al.*, 2014; Kuczek & Rogers, 1987)
	KAP19	7+4	(Rogers *et al.*, 2004b)		
	KAP20	2	(Rogers *et al.*, 2004b)		
	KAP21	2+1	(Rogers *et al.*, 2004b)		
	KAP22	1	(Rogers *et al.*, 2004b)		
		88+17		**23+2**	

[†] Numbers after "+" represent pseudogenes.

[φ] Membership is based on the identification of the gene and not the protein.

The known sheep genes are clustered into three domains on three chromosomes (Table 3-2). Given the similarities in the chemical make-up and structure of wool and human hair, and the similarity of the individual genes or clustered families (Table 3-2), it is expected that more individual *KRTAPs* and families will be found in the sheep genome. Equally it would not be unreasonable to expect that the overall number of putative *KRTAPs* might be higher than previously thought. While the discovery of more and more KAP genes is not in itself any reason to change the nomenclature system, it does highlight the need to revise its suitability, especially in light of the issues described below.

Table 3-2. Chromosomal locations of the KAP genes in humans and sheep

Species	Chromosomal location	KAP gene families
Human	17q21.2	KAP1 - KAP4, KAP9, KAP16, KAP17
Sheep	11	KAP1 - KAP4
Human	21q22.3	KAP10, KAP12
Sheep	unknown	Un-identified
Human	11p15.5 and 11q13.4	KAP5
Sheep	21	KAP5
Human	21q22.1	KAP6 - KAP8, KAP11, KAP13, KAP15, KAP19 - KAP22, KAP24 - KAP27
Sheep	1	KAP6 - KAP8, KAP11, KAP13, KAP24

3.3.1 Impact of Species of Origin on KAP/*KRTAP* Classification

The species of origin of any KAP protein or gene sequence may need to be stated in any classification system. Homology comparisons between species reveals that KAP proteins of the same name (and therefore one would assume family), may actually have low inter-species homology, but never-the-less tend to cluster with other family members from the same species. For example, sheep KAP1-3 does not share a high homology with KAP1-3 from either human, mouse or dog, but is more similar to other sheep KAP1 family members (Figure 3-1A). All the sheep KAP1 family members tend to cluster, and all the human KAP1 family members tend to cluster, on sequence-based comparison (Figure 3-1A). A similar phenomenon can also be seen for the KAP3, KAP4 and KAP5 families (Figure 3-1B, C and D). Consequently, it can be difficult to assign a new sequence to a family, or constituent group within that family, especially if that sequence is one of the first obtained from a species.

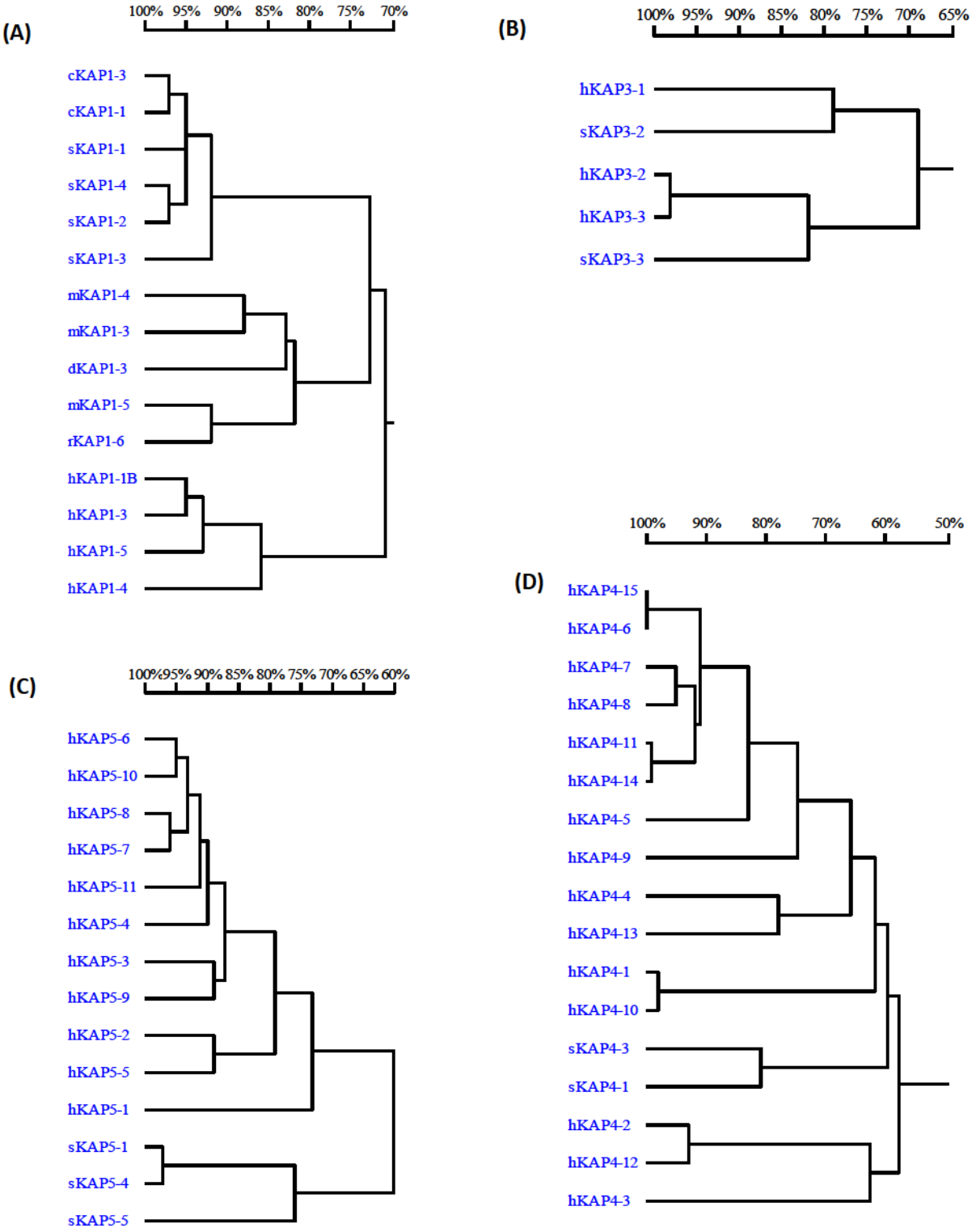


Figure 3-1. Phylogenetic relationships of KAP family members from different species. Phylogenetic trees are constructed using DNAMAN based on the predicated amino acid sequences for KAP1 (A), KAP3 (B), KAP5 (C) and KAP4 (D) families which have two or more family members identified in human and sheep. The scales show% amino acid match between individual proteins or branches. Prefixes h, s, c, d, m and r represent human, sheep, cattle, dog, mouse and rat, respectively.

The assignment of KAP/*KRTAP* family membership is already in some instances not supported by homology (Figure 3-1), and some assignments seem strange. For example, Liu *et al.* (2009) recently reported the presence of a *Capra hircus* KAP16-6 gene, although it has very low sequence homology with any human KAP genes, including the published human KAP16-1 gene sequence [GenBank AC003958; (Rogers *et al.*, 2001)], and in the absence of evidence of there being a KAP16-2, KAP16-3, KAP16-4 or KAP16-5 gene in humans. Hence it is believed that at the very least, the species of origin of any given sequence needs to be clearly identified, although the need to identify species might be less important in literature if a publication is focussed on a single species, and no conclusions that may have implications across species are drawn.

Researchers finding new KAP or *KRTAP* sequences probably also need to be more cautious in assigning the sequence to a family. Accordingly, it is recommended that the "L" term for "like", is used if any doubt exists at all as to the origin of the sequence, and that this term, which is also used in Powell and Rogers' (1997) system, continues in the nomenclature recommended below.

3.3.2 Genetic Variation in the *KRTAPs*

Studies of variation in *KRTAPs* are limited in most species, and knowledge of the various types of genetic variation is therefore also limited. Currently humans and sheep are the two most studied species, probably because of the recent emphasis on human genome discovery and the historic importance of wool in making textiles. It is expected however that as more genomes are sequenced in more individuals, the need for a comprehensive nomenclature that effectively accommodates genetic variation will increase.

In humans, studies of *KRTAP* variation have only been carried out in Caucasian and Japanese individuals and have been restricted to the *KRTAP1* (Shimomura *et al.*, 2002) and *KRTAP4* (Kariya *et al.*, 2005) families. Four previously identified and apparently different KAP1-n genes (*KRTAP1-1A, KRTAP1-1B, KRTAP1-6* and *KRTAP1-7*) have been shown to be allelic variants of a single gene (Shimomura *et al.*, 2002), while four other KAP1-n genes (*KRTAP1-8A, KRTAP1-8B, KRTAP1-3* and *KRTAP1-9*) were revealed to be allelic variants of another KAP1-n gene (Shimomura *et al.*, 2002). Two or three allelic variants have also been reported for 10 of the 11 human KAP4-n genes (Kariya *et al.*, 2005).

In sheep, genetic variation has been reported for the *KRTAP1* (Gong *et al.*, 2010a; Gong *et al.*, 2011e; Itenge-Mweza *et al.*, 2007; Rogers *et al.*, 1994b), *KRTAP3* (McLaren *et al.*, 1997), *KRTAP5* (Gong *et al.*, 2010c; McLaren *et al.*, 1997), *KRTAP6* (Gong *et al.*, 2011a), *KRTAP7* (Gong *et al.*, 2011d; McLaren *et al.*, 1997) *KRTAP8* (Gong *et al.*, 2014; Gong *et al.*, 2011d; Wood *et al.*, 1992), *KRTAP11* (Gong *et al.*, 2011b), *KRTAP13* (Gong *et al.*, 2011c) and *KRTAP24* (Zhou *et al.*, 2012) families. Up to nine alleles

have been reported for *KRTAP1-2* (Gong *et al.*, 2011e), *KRTAP1-3* (Itenge-Mweza *et al.*, 2007) and *KRTAP1-4* (Gong *et al.*, 2010a). It should be noted that the apparently higher degree of variation found in sheep, compared to humans, is possibly due to a greater number of genomes being screened.

The variation detected in the KAP genes includes single nucleotide substitutions and length variation. Variation in length is noted in both sheep and human KAP1-n genes (Rogers *et al.*, 1994b; Shimomura *et al.*, 2002), and human KAP4-n genes (Kariya *et al.*, 2005). It appears to be the result of having a variable number of cysteine-rich repeated coding sequences and these have probably arisen by intragenic deletion and/or duplication of repeated segments of the genes during evolution (Kariya *et al.*, 2005).

There is little understanding of how variation in human KAP genes affects hair structure or other keratinagous tissue. However, genetic variation in each individual KAP family could be much higher than previously thought based on recent research in sheep (Gong *et al.*, 2010a; Gong *et al.*, 2011a; Gong *et al.*, 2011b; Gong *et al.*, 2010c; Gong *et al.*, 2011c; Gong *et al.*, 2011d; Gong *et al.*, 2011e; Zhou *et al.*, 2012), and this may underpin some of the variation in hair and wool characteristics. The ability of the nomenclature system to accommodate extensive genetic variation, or what more correctly might be called polymorphism, given the over-use of this word in describing less variable gene systems; is one key driver in this revised system. What is more, while this revised nomenclature is not dissimilar to that of Powell and Rogers (1997) in accommodating genetic variation, it is felt that the term in the name that denotes genetic or allelic variation should be at the end of the gene name, especially as we believe an increased emphasis on how this variation affects wool and hair traits will emerge with time.

3.4 The Revised Nomenclature for Mammalian KAPs/*KRTAPs*

Given the numerous KAP genes identified to date and the high levels of diversity among the KAPs from different species of more distant phylogenetic relationship, a revised nomenclature needed to be both flexible in accommodating variation and informative. It is proposed that the current KAP nomenclature (Powell and Rogers, 1997) could be easily modified to accommodate the HGNC guidelines to become a consensus system for all mammalian species, as follows:

sp-KAPm-nL*x = protein

sp-*KRTAPm-n(p/L)*x* = gene

In this nomenclature, “sp” is a unique letter-based code for different species described by the protein knowledge-based UniProt (www.uniprot.org/docs/speclist), for example, “HUMAN” for *Homo sapiens*, “SHEEP ” for *Ovis aries*, “BOVIN” for *Bos taurus* and “MOUSE” for *mus musculus*. This would typically only be used in publications and when necessary; “m” is a number identifying the family; “n” is a constituent of that family; “p” signifies a pseudogene if there is an obvious fault in the gene (e.g. presence of an unexpected stop codon), while “L” if present, signifies “like” and refers to a temporary “place-holder” until the family (or constituency within that family) is confirmed; “x” is an alphabetical letter signifying the variant or allele, but preferably at the level of an extended haplotype encompassing the promoter, 5’-UTR, exons and introns and 3’UTR. “p” and “L” should probably not be used together as a pseudogene will not produce a protein by definition, and hence genetic homology with any other sequence is of limited value.

The use of KAP for the protein and *KRTAP* for the gene was consistent with the keratin nomenclature where K is used for a protein and *KRT* is used for the associated gene.

To allow adequate time for transition and capture historically useful information, it is suggested that in future, historic names might be bracketed after the first mention of the gene or protein, such as *KRTAP1-1*(*B2A*), for the next few years [see old terminology in Powell and Rogers (1997)].

3.5 Concluding Remarks

This nomenclature preserves the widely used and broadly referred to system proposed by Powell and Rogers (1997), but with the important addition of a species identifier to constitute an informative KAP naming system and a minor change in the order of terms in the nomenclature. Accordingly, the new nomenclature should have minimal impact on current publications and databases, but if used correctly should facilitate more informed discussion about the KAP genes and proteins. It is strongly urged authors to use the L-term until such time as the location of the gene can be confirmed on the chromosome and its similarity to a sequence of known classification confirmed, or if they harbour any concerns about the family of origin of the protein or gene they are describing. It may be appropriate for scientists working on KAP genes and proteins to communicate more freely with each other and also with organisations like HGNC to insure that the nomenclature is used appropriately, while also being regularly revised to accommodate any future findings about the KAP genes and proteins.

Chapter 4

Association of Variation in Two KAP Genes and Wool Traits

This work has been published in "Gong H, Zhou H, Hodge S, Dyer JM and Hickford JGH. (2015). Association of wool traits with variation in the ovine KAP1-2 gene in Merino cross lambs. Small Ruminant Research 124:24-9", and "Zhou H, Gong H*, Li S, Luo Y and Hickford JGH. (2015). A 57-bp deletion in the ovine KAP6-1 gene affects wool fibre diameter. Journal of Animal Breeding and Genetics 132:301-7 (*co-first authors)".*

4.1 Introduction

The results of Chapter 2 suggest that the extent of KAP diversity in sheep possibly higher than that found in humans. Furthermore, all the KAP gene investigated are polymorphic in sheep. As outlined in Section 1.5, KAP genes have been implicated in imparting particular wool characteristics. Of the ovine KAP gene families identified, two gene families, KAP1 and KAP6, exhibit some interesting characteristics making them as prime candidate genes for the development of genetic markers.

The KAP1 genes are located on ovine chromosome 11. In this chromosome region, a number of QTLs for wool traits, including MSS, CVFD, GFW and CFW, have been reported (Rogers *et al.*, 1994a; Roldan *et al.*, 2010). Given that the KAP1 genes are highly variable (Gong *et al.*, 2010a; Gong *et al.*, 2011e; Itenge-Mweza *et al.*, 2007) and are expressed early in the cortical cells where many *KRTs* are expressed (soon after the expression of the HGT-KAP genes) (Powell & Rogers, 1997; Yu *et al.*, 2009), variation in KAP1 genes may affect wool traits. A representative member of the KAP1 family; the KAP1-2 gene (*KRTAP1-2*) will be investigated in this chapter to ascertain whether it is associated with variation in wool traits.

While the composition of KAP proteins from each group varies spatially and developmentally within the wool fibre (Powell & Rogers, 1990), the most significant variation is seen within the HGT group where their abundance varies considerably, ranging from less than 1% in Lincoln sheep wool, to between 4% and 12% in Merino wool (Gillespie, 1990). HGT-KAPs are predominantly present in the cortex (Powell & Rogers, 1997) and a reduction in the content of HGT-KAPs appears, at least in part, to be responsible for the felting lustre mutant found in Merino sheep (Li *et al.*, 2009). The wide range in proportion of HGT-KAPs in different wools raises intriguing questions about the function of these proteins in the fibre.

There are three HGT-KAP families which have been identified in sheep: KAP6, KAP7 and KAP8 (Gong *et al.*, 2012). Of these, KAP6 is the most diverse HGT-KAP family, containing at least three variable KAP6 genes (Gong *et al.*, 2011a). The high level of variation appears to be observed in *KRTAP6-1* which not only exhibits sequence variation but also length variation (Gong *et al.*, 2011a).

These HGT-KAP gene families are clustered on ovine chromosome 1 (McLaren *et al.*, 1997). Parsons *et al.* (1994b) reported that variation in a KAP6 gene was linked to wool fibre diameter, but the association was only observed in one of the eight Merino sire groups. The variation was detected using the now superseded RFLP-Southern hybridisation technique, with no detailed sequence information being available for the variation that was detected. Given the potential effect of the KAP6 genes on the wool fibre characteristics, further investigation is needed to ascertain whether variation in any of the KAP6 genes affects wool traits. The potential effect of variation in *KRTAP6-1* on wool traits was also investigated in this chapter.

4.2 Materials and Methods

4.2.1 Sheep Blood and Wool Samples

A total of 383 Merino × Southdown-cross lambs, sired by six rams, were investigated in this study. Ram lambs were castrated at approximately one month of age to limit the influence of male hormones on wool characteristics. A blood sample from each lamb was collected onto an FTA card and genomic DNA for PCR amplification was purified from the dried blood spots using a two-step procedure (Zhou *et al.*, 2006).

All the lambs were shorn at 12 months of age and a wool sample from the mid-side region of each lamb was collected. GFW (kg) was recorded at shearing, and other wool traits were measured by the New Zealand Wool Testing Authority Ltd (Napier, New Zealand), including CFW (kg), MFD (μm), FDSD (μm), CVFD (%), MSL (mm), MFC (°/mm), MSS (N/ktex), and PF (%). Wool yield (%) was calculated as the proportion of CFW, relative to GFW.

4.2.2 PCR-SSCP Genotyping

Based on the sequences of nine known ovine *KRTAP1-2* variants (Gong *et al.*, 2011e), and the comparison with previously reported *KRTAP1-1*, *KRTAP1-3* and *KRTAP1-4* sequences (Gong *et al.*, 2010a; Itenge-Mweza *et al.*, 2007), two polymerase chain reaction (PCR) primers, 5'-^{375}CCAGCCAACTTCCATCCAA^{393}-3' and 5'-^{762}AAATCAGCAGCCTTGCTTTC^{743}-3' (nucleotide coordinates relative to the *KRTAP1-2* sequence HQ897973), were designed to amplify a small (388 bp) variable region of *KRTAP1-2* that would be suitable for SSCP analysis.

Based on the published ovine *KRTAP6* sequences (Fratini *et al.*, 1993; Gong *et al.*, 2011a) and the Oar_v3.1 chromosome 1 sequence (NC_019458), two PCR primers, 5'-^{1011T}CTACCCGAGAACAACCTC^{1030}-3' and 5'-^{1390}AGGCAAGTCTTTAGTAGGAC^{1373} -3' (nucleotide coordinates relative to the *KRTAP6-1* sequence M95719), were designed to specifically amplify the entire coding sequence of *KRTAP6-1*.

PCR amplification was performed in a 15 µL reaction containing the genomic DNA on one 1.2 mm punch of FTA paper, 0.25 µM of each primer, 150 µM of each dNTP (Bioline, London, UK), 2.5 mM of Mg^{2+}, 0.5 U of *Taq* DNA polymerase (Qiagen, Hilden, Germany) and 1× reaction buffer supplied with the enzyme. The thermal profile consisted of 2 min at 94 °C, followed by 35 cycles of 30 s at 94 °C, 30 s at 60 °C and 30 s at 72 °C, with a final extension of 5 min at 72 °C. Amplification was carried out in a S1000 thermal cycler (Bio-Rad, Hercules, CA, USA).

PCR amplicons were subject to SSCP analysis. SSCP analysis was carried out in 14% polyacrylamide gels using 0.5 × TBE buffer at either 34 °C 200 V for 20 h (for *KRTAP1-2*) or 15 °C 340 V for 18 h (for *KRTAP6-1*). For each gene, amplicons of known ovine *KRTAP1-2* (Gong *et al.*, 2011e) or *KRTAP6-1* (Gong *et al.*, 2011a) variants were included in each gel as reference standards, and based on the SSCP banding patterns the variants present in individual samples were determined. The gels were silver-stained according to the method of Byun *et al.* (2009).

Amplicons showing new SSCP patterns were subsequently sequenced using a rapid approach described previously (Gong *et al.*, 2011a). Briefly, a band corresponding to the variant was excised as a gel slice from the polyacrylamide gel, macerated, and then used as a template for re-amplification with the original primers. This second amplicon was then directly sequenced. Each PCR amplicon was sequenced in both directions until an identical sequence was obtained from at least two independent PCR reactions. Sequencing was carried out at the Lincoln University DNA sequencing Facility.

4.2.3 Statistical Analyses

Statistical analyses were performed using Minitab version 16. Firstly, Pearson correlation coefficients were calculated to test the strength of the relationship between the ten wool traits: GFW, CFW, Yield, MFD, FDSD, CVFD, MSL, MSS, MFC and PF. Next, General Linear Models (GLMs) were used to assess the effect of the presence/absence of the *KRTAP1-2* or *KRTAP6-1* variants on these wool traits. Initially, 'single-variant' models were performed to ascertain which variants would be included in a second series of 'multi-variant' models, where any variant that had an association with a trait in the 'single-variant' models with $P < 0.2$, and which could thus potentially impact on the wool trait being tested, were included. A third series of GLMs were used to compare various wool traits among lambs that had different genotypes, but only if the genotype frequency was greater than 5%, and thereby

representing an adequate sample size. When the GLMs indicated significant differences among genotypes, multiple pairwise comparisons were made using a Tukey test to correct for family-wise error rate.

In all models, gender and birth rank were fitted as fixed factors, and sire was fitted as a random factor. Only main effects were tested.

4.3 Results

4.3.1 Correlations between Wool Traits

Strong correlations ($|r| > 0.7$) were found between GFW and CFW; between MFD and FDSD; between CVFD and FDSD; and between PF, and MFD and FDSD (Table 4-1). Moderate correlations ($0.3 < |r| \leq 0.7$) were found between GFW, and MSL and MFC; between CFW, and yield, MSL and MSS; between yield, and MFD and MFC; between MFD and MFC; between MSS, and FDSD and CVFD; between MSL and MFC; and between CVFD and PF (Table 4-1). All these correlations were highly significant ($P < 0.001$). There were only weak or negligible correlations ($|r| \leq 0.3$) between other wool traits (Table 4-1). For example, MFD was not correlated with either GFW or CFW ($P > 0.05$).

Table 4-1. Pearson correlation coefficients between various wool traits for 368 Merino-cross sheep[1]

Trait[2]	GFW	CFW	Yield	MFD	FDSD	CVFD	MSL	MSS	MFC
CFW	**0.920*****								
Yield	0.206***	0.558***							
MFD	0.065	-0.074	-0.357***						
FDSD	-0.081	-0.157**	-0.262***	**0.733*****					
CVFD	-0.173**	-0.167**	-0.088	0.304***	**0.813*****				
MSL	0.400***	0.435***	0.241***	-0.054	-0.072	-0.058			
MSS	0.186***	0.256***	0.282***	-0.163**	-0.325***	-0.336***	-0.068		
MFC	-0.400***	-0.533***	-0.518***	0.238***	0.187***	0.073	-0.386***	-0.206***	
PF	0.080	-0.009	-0.203***	**0.788*****	**0.778*****	0.416***	-0.053	-0.166**	0.076

[1] Correlations with $|r| > 0.7$ are in bold, and those with $0.3 < |r| \leq 0.7$ are underlined. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

[2] GFW – Greasy Fleece Weight; CFW – Clean Fleece Weight; MFD – Mean Fibre Diameter; FDSD – Fibre Diameter Standard Deviation; CVFD – Coefficient of Variation of Fibre Diameter; MSL – Mean Staple Length; MSS – Mean Staple Strength; MFC – Mean Fibre Curvature; PF – Prickle Factor (percentage of fibres over 30 microns).

4.3.2 Variation in *KRTAP1-2* and its Association with Wool Traits

Allele variants and genotypes detected by PCR-SSCP

Under the PCR-SSCP conditions developed for ovine *KRTAP1-2*, all known variants (*A* - *I*) (Gong *et al.*, 2011e) exhibited unique PCR-SSCP banding patterns (Figure 4-1), and the variants present in the samples could be determined. Of the nine previously reported variants, seven (*A* - *C*, *E* - *H*) were detected in this Merino × Southdown-cross flock. Two new variants (*J* and *K*) were also identified. Sequence analysis revealed that these two new variants resulted from the combination of previously identified SNPs, rather than new SNPs. The sequences of variants *J* and *K* have been deposited into the GenBank with accession numbers KM105941 and KM105942.

The frequencies of the *KRTAP1-2* variants detected in these sheep were: *A*: 25.6%; *B*: 10.4%; *C*: 1.5%; *E*: 0.3%; *F*: 39.5%; *G*: 12.9%; *H*: 9.1%; *J*: 0.4%; and *K*: 0.3%. There were 23 different genotypes observed. Of these, only six occurred at a frequency over 5%, and these were *AF*: 27.2%; *AG*: 6.3%; *BF*: 11.2%; *FF*: 8.4%; *FG*: 12.5%; and *FH*: 9.1%. The remaining 17 genotypes were present at frequencies under 5%, including *AA, AB, AC, AH, BB, BC, BG, BH, CF, CG, EF, FJ, FK, GG, GH, GJ* and *HH*.

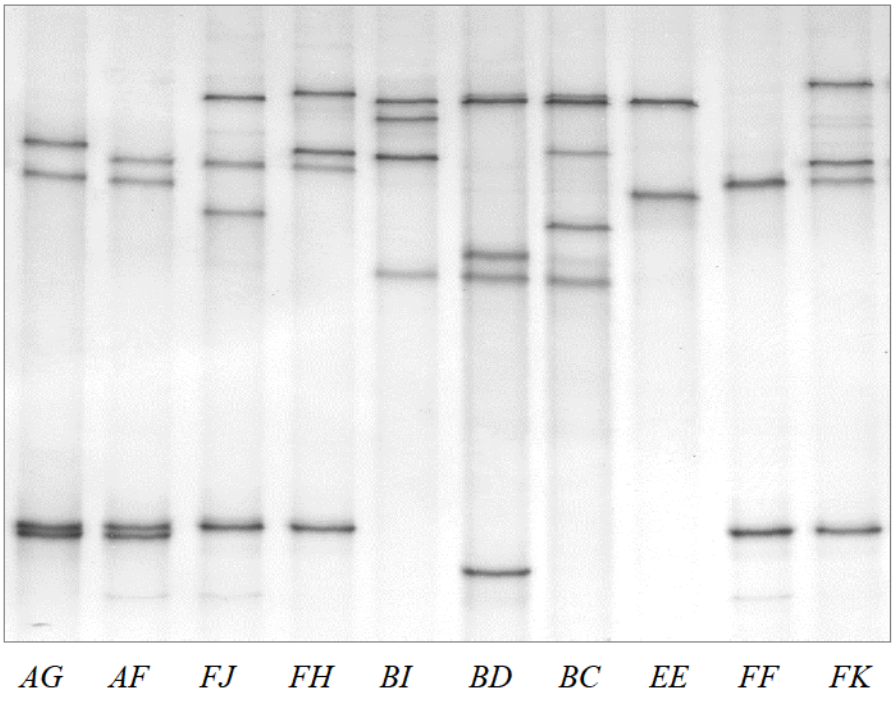


Figure 4-1. PCR-SSCP of ovine *KRTAP1-2*. Eleven PCR-SSCP patterns corresponding to nine previously (*A* - *I*) and two newly (*J* and *K*) identified variant sequences are shown.

Associations between variation in *KRTAP1-2* and variation in wool traits

Of the nine variants detected in these lambs, three variants (*E, J* and *K*) occurred at a very low frequency (< 0.5%), and accordingly their associations with wool traits were not tested. The presence of *A* was found to be associated with an increase in both GFW and CFW, and a decrease in MFC in the 'single-variant' models (Table 4-2). These effects persisted when corrected for other variants in the 'multi-variant' models. In the 'single-variant' model, the presence of *A* was not associated with wool yield, but was found to be associated with increased wool yield in the 'multi-variant' model (Table 4-2).

The presence of *B* was associated with an increase in CFW, yield and MSL in the 'single-variant' models, and these associations persisted in the multi-variant models (Table 4-2). In the 'single-variant' model, the presence of B was not associated with GFW, but was found to be associated with increased GFW in the 'multi-variant' model (Table 4-2).

In 'single-variant' models, the presence of *C* was found to be associated with a decrease in GFW, CFW, yield and MSL; and an increase in FDSD and CVFD (Table 4-2). These associations persisted when corrected for the other variants. Variant *C* did not show any effect on PF in the 'single-variant' model, but its presence was associated with an increase in PF in the 'multi-variant' model (Table 4-2).

No association with wool traits was found for the other variants except for the association between variant H and MSS in the 'multi-variant' model (Table 4-2).

Effect of common *KRTAP1-2* genotypes on wool traits

With the six common *KRTAP1-2* genotypes (*AF, AG, BF, FF, FG* and *FH*), an effect of genotype was observed for GFW, CFW and yield (Table 4-3). Sheep of genotypes *AF* and *AG* produced a GFW approximately 9% higher than those of genotypes *FG* and *FH*. Genotypes *AF* and *AG* had CFW approximately 13% higher than *FG*. In terms of yield, *AF, AG* and *BF* had a high mean yield compared to *FF* and *FG* (Table 4-3).

4.3.3 Variation in *KRTAP6-1* and its Association with Wool Traits

Variation in ovine *KRTAP6-1*

There were three PCR-SSCP banding patterns detected for ovine *KRTAP6-1*, with either one or a combination of two banding patterns observed in each sheep (Figure 4-2). DNA sequencing revealed that these three PCR-SSCP patterns represented three different nucleotide sequences (*A, B* and *C*).

Table 4-1. Association of *KRTAP1-2* variants with various wool traits (mean ± SE)[1]

Trait[2]	Variant	n		Single-variant model			Multi-variant model			
		Absent	Present	Absent	Present	*P*	Other[3]	Absent	Present	*P*
GFW	*A*	205	178	2.43 ± 0.04	2.53 ± 0.04	**0.037**	*B,C,F*	2.09 ± 0.06	2.22 ± 0.07	**0.009**
(kg)	*B*	304	79	2.46 ± 0.03	2.55 ± 0.05	*0.086*	*A,C,F*	2.09 ± 0.06	2.62 ± 0.07	**0.013**
	C	372	11	2.49 ± 0.03	1.72 ± 0.11	**<0.001**	*A,B,F*	2.57 ± 0.05	2.04 ± 0.16	**<0.001**
	F	112	271	2.41 ± 0.05	2.51 ± 0.03	0.143	*A,B,C*	2.13 ± 0.06	2.17 ± 0.07	0.520
	G	292	91	2.48 ± 0.03	2.48 ± 0.05	0.881				
	H	311	72	2.48 ± 0.03	2.45 ± 0.05	0.475				
CFW	*A*	205	178	1.74 ± 0.03	1.84 ± 0.03	**0.017**	*B,C*	1.54 ± 0.05	1.67 ± 0.05	**0.002**
(kg)	*B*	304	79	1.77 ± 0.02	1.87 ± 0.04	**0.015**	*A,C*	1.54 ± 0.05	1.68 ± 0.06	**0.001**
	C	372	11	1.80 ± 0.02	1.37 ± 0.09	**<0.001**	*A,B*	1.84 ± 0.03	1.38 ± 0.09	**<0.001**
	F	112	271	1.75 ± 0.04	1.81 ± 0.03	0.292				
	G	292	91	1.79 ± 0.02	1.78 ± 0.04	0.682				
	H	311	72	1.80 ± 0.02	1.76 ± 0.04	0.357				
Yield	*A*	205	178	71.8 ± 0.54	72.8 ± 0.57	0.162	*B,C,F*	74.4 ± 0.94	75.9 ± 1.02	**0.040**
(%)	*B*	304	79	71.9 ± 0.44	73.7 ± 0.74	**0.017**	*A,C,F*	74.0 ± 0.89	76.2 ± 1.08	**0.006**
	C	372	11	72.2 ± 0.38	76.7 ± 1.67	**0.007**	*A,B,F*	72.9 ± 0.46	77.4 ± 1.69	**0.008**
	F	112	271	73.2 ± 0.74	71.9 ± 0.45	0.167	*A,B,C*	75.2 ± 0.99	75.1 ± 1.05	0.950
	G	292	91	72.5 ± 0.45	71.7 ± 0.68	0.333				
	H	311	72	72.3 ± 0.44	71.5 ± 0.79	0.352				
MFD	*A*	205	178	19.6 ± 0.16	19.5 ± 0.17	0.605				
(µm)	*B*	304	79	19.6 ± 0.13	19.6 ± 0.22	0.851				
	C	372	11	19.5 ± 0.12	19.8 ± 0.51	0.629				
	F	112	271	19.2 ± 0.22	19.7 ± 0.15	*0.073*				
	G	292	91	19.6 ± 0.13	19.5 ± 0.24	0.525				
	H	311	72	19.5 ± 0.12	19.8 ± 0.23	0.357				
FDSD	*A*	205	178	4.16 ± 0.06	4.18 ±0.07	0.848				
(µm)	*B*	304	79	4.19 ± 0.05	4.10 ± 0.09	0.372				
	C	372	11	4.17 ± 0.05	4.84 ± 0.20	**0.005**				
	F	112	271	4.09 ± 0.09	4.22 ± 0.06	0.279				
	G	292	91	4.19 ± 0.05	4.17 ± 0.09	0.839				
	H	311	72	4.18 ± 0.05	4.19 ± 0.10	0.960				
CVFD	*A*	205	178	21.2 ± 0.26	21.4 ± 0.28	0.579				
(%)	*B*	304	79	21.3 ± 0.21	21.0 ± 0.35	0.396				
	C	372	11	21.3 ± 0.17	23.6 ± 0.77	**0.002**				
	F	112	271	21.3 ± 0.34	21.3 ± 0.22	0.973				
	G	292	91	21.3 ± 0.21	21.4 ± 0.38	0.852				
	H	311	72	21.3 ± 0.19	21.1 ± 0.34	0.458				
MSL	*A*	205	178	79.1 ± 1.05	80.6 ± 1.11	0.328				
(mm)	*B*	304	79	78.9 ± 0.86	82.9 ± 1.43	**0.009**	*C,G*	75.2 ± 1.81	78.8 ± 2.30	**0.028**
	C	372	11	79.8 ± 0.77	72.8 ± 3.32	**0.032**	*B,G*	80.7 ± 1.00	73.3 ± 3.43	**0.027**
	F	112	271	80.3 ± 1.69	79.2 ± 0.97	0.601				
	G	292	91	80.0 ± 0.86	77.6 ± 1.51	0.158	*B,C*	77.5 ± 1.79	76.4 ± 2.33	0.495
	H	311	72	79.7 ± 0.84	78.5 ± 1.53	0.477				
MSS	*A*	205	178	26.3 ± 0.77	25.7 ± 0.81	0.552				
(N/ktex)	*B*	304	79	25.7 ± 0.64	27.3 ± 1.06	0.168	*H*	26.4 ± 0.70	28.1 ± 1.13	0.123
	C	372	11	26.1 ± 0.57	23.9 ± 2.51	0.381				
	F	112	271	27.1 ± 1.10	25.6 ± 0.71	0.272				
	G	292	91	26.1 ± 0.64	25.8 ± 1.00	0.767				
	H	311	72	25.6 ± 0.60	27.8 ± 1.10	*0.063*	*B*	26.1 ± 0.67	28.4 ± 1.16	**0.048**
MFC	*A*	205	178	93.7 ±1.32	90.0 ± 1.38	**0.036**				
(°/mm)	*B*	304	79	91.8 ± 1.09	92.9 ± 1.80	0.578				
	C	372	11	92.0 ± 0.93	90.6 ± 4.10	0.714				
	F	112	271	91.4 ± 2.12	92.3 ± 1.22	0.709				
	G	292	91	91.6 ± 1.08	93.6 ± 1.90	0.356				
	H	311	72	92.0 ± 1.05	92.4 ± 1.92	0.850				
PF	*A*	205	178	2.82 ± 0.30	2.86 ± 0.32	0.923				
(%)	*B*	304	79	2.95 ± 0.26	2.75 ± 0.43	0.748				
	C	372	11	2.80 ± 0.22	4.47 ± 0.94	*0.076*	*F*	2.60 ± 0.24	4.70 ± 0.94	**0.029**
	F	112	271	2.33 ± 0.41	3.06 ± 0.26	0.152	*C*	3.15 ± 0.55	4.16 ± 0.56	*0.055*
	G	292	91	2.88 ± 0.26	2.98 ± 0.45	0.841				
	H	311	72	2.97 ± 0.25	2.60 ± 0.45	0.440				

[1] Predicted means, standard errors and *P* values from General Linear Models (GLMs). $P \leq 0.05$ are in bold whereas $0.05 < P < 0.1$ are in italics.

[2] GFW – Greasy Fleece Weight; CFW – Clean Fleece Weight; MFD – Mean Fibre Diameter; FDSD – Fibre Diameter Standard Deviation; CVFD – Coefficient of Variation of Fibre Diameter; MSL – Mean Staple Length; MSS – Mean Staple Strength; MFC – Mean Fibre Curvature; PF – Prickle Factor (percentage of fibres over 30 microns).

[3] Other variants fitted into the models.

Table 4-3. The effect of *KRTAP1-2* genotype on various wool traits

Trait[1]	Mean ± SE[2]						*P*
	AF (n=104)	*AG* (n=24)	*BF* (n=43)	*FF* (n=32)	*FG* (n=48)	*FH* (n=35)	
GFW (kg)	2.51 ± 0.04^{a}	2.52 ± 0.07^{ab}	2.43 ± 0.05^{ab}	2.36 ± 0.06^{ab}	2.29 ± 0.05^{b}	2.29 ± 0.06^{b}	**0.003**
CFW (kg)	1.82 ± 0.03^{a}	1.85 ± 0.06^{a}	1.77 ± 0.05^{ab}	1.64 ± 0.05^{ab}	1.61 ± 0.04^{b}	1.65 ± 0.05^{ab}	**<0.001**
Yield (%)	72.6 ± 0.65^{ab}	73.8 ± 1.21^{a}	73.2 ± 0.89^{a}	69.6 ± 1.06^{b}	70.4 ± 0.87^{b}	71.3 ± 1.01^{ab}	**0.019**
MFD (µm)	19.4 ± 0.20	18.9 ± 0.36	19.4 ± 0.27	19.4 ± 0.32	19.1 ± 0.26	19.7 ± 0.30	0.510
FDSD (µm)	4.21 ± 0.08	4.18 ± 0.13	4.17 ± 0.12	4.12 ± 0.13	4.19 ± 0.12	4.22 ± 0.12	0.658
CVFD (%)	21.3 ± 0.30	22.1 ± 0.52	21.4 ± 0.44	21.4 ± 0.50	21.5 ± 0.43	21.6 ± 0.45	0.852
MSL (mm)	81.6 ± 1.28	82.0 ± 2.35	84.1 ± 1.73	79.3 ± 2.05	78.1 ± 1.69	80.1 ± 1.96	0.146
MSS (N/ktex)	24.5 ± 0.85	25.3 ± 1.56	26.0 ± 1.15	23.9 ± 1.36	24.1 ± 1.13	26.3 ± 1.31	0.575
MFC (°/mm)	90.0 ± 1.52	85.6 ± 2.79	92.9 ± 2.05	94.9 ± 2.43	95.0 ± 2.00	92.0 ± 2.33	0.102
PF (%)	2.83 ± 0.39	1.88 ± 0.71	2.06 ± 0.53	2.17 ± 0.62	2.20 ± 0.52	3.12 ± 0.59	0.543

[1] GFW – Greasy Fleece Weight; CFW – Clean Fleece Weight; MFD – Mean Fibre Diameter; FDSD – Fibre Diameter Standard Deviation; CVFD – Coefficient of Variation of Fibre Diameter; MSL – Mean Staple Length; MSS – Mean Staple Strength; MFC – Mean Fibre Curvature; PF – Prickle Factor (percentage of fibres over 30 microns).

[2] Estimated marginal means, standard errors and *P* values derived from GLMs. Bonferroni correction fitted for repetitive testing. $P \leq 0.05$ are in bold. Means within rows that do not share a superscript letter are different ($P < 0.05$).

Sequences *A* and *B* had identical coding sequence, with only three nucleotide differences occurring downstream of the stop codon. However, sequence *C* had a deletion of a 57 bp sequence within the coding sequence. This 57 bp deletion would notionally result in the loss of 19 amino acids in the central region of the protein (Figure 4-3).

The *B* and *C* sequences completely matched with two of this study's previously identified *KRTAP6* sequences GU319875 and GU319873, respectively (Gong *et al.*, 2011a), whereas sequence *A* was unique and represented a newly identified variant of ovine *KRTAP6-1*.

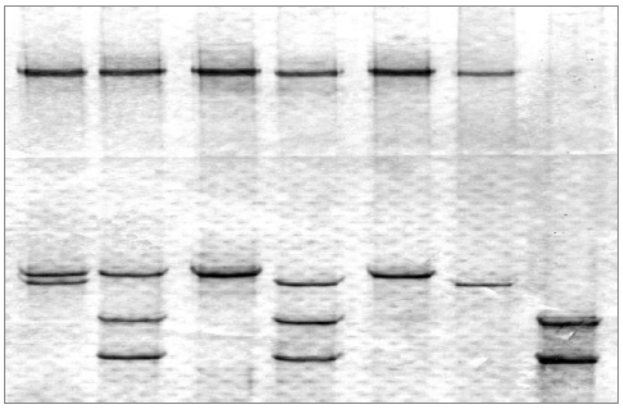


Figure 4-2. PCR-SSCP of ovine *KRTAP6-1*. Three banding patterns representing three variants (*A, B* and *C*) were identified in both homozygous and heterozygous forms.

```
KAP6-1*A    MCGYYGNYYGGLGCGSYSYGGLGCGYGSCYGSGFRRLGCGY      41
KAP6-1*B    ---------------------------------------      41
KAP6-1*C    ----------------------.............      26

KAP6-1*A    GCGYGYGSRSLCGSGYGYGSRSLCGSGYGCGSGYGSGFGYYY    83
KAP6-1*B    ---------------------------------------      83
KAP6-1*C    ....-----------------------------------      64
```

Figure 4-3. Alignment of predicted amino acid sequences for three ovine *KRTAP6-1* variants. Amino acid residues identical to the top sequence are indicated by dashes, and dots represent deletions. Aromatic amino acids are shown in red, whereas basic amino acids are shown in green.

Effect of variation in *KRTAP6-1* on wool traits

Six genotypes were detected among these sheep, and they were: *AA* (22.8%), *AB* (40.5%), *AC* (4.1%), *BB* (26.1%), *BC* (5.4%) and *CC* (1.1%). This gave variant frequencies of 45.1%, 49.1% and 5.8% for *A, B* and *C* respectively. There were more *AA*, *BB* or *CC* but less *AB*, *AC* or *BC* than expected, and the *P* value of the chi-square test for Hardy-Weinberg Equilibrium (HWE) was 0.02.

The presence of variant *C* was found to be associated with an increase in MFD, FDSD, CVFD and PF, and decreased yield in the 'single-variant' model, and these associations persisted in the 'multi-variant' model. The increase was 8.6%, 13.3%, 4.8% and 180.4% respectively for MFD, FDSD, CVFD and PF, and the decrease in yield was 4.4% (Table 4-4).

Table 4-4. Association of *KRTAP6-1* variants with various wool traits (mean ± SE)[1]

Trait[2]	Variant	n		Single-variant model			Multi-variant model			
		Absent	Present	Absent	Present	*P*	Other variant fitted	Absent	Present	*P*
GFW	*A*	120	248	2.22 ± 0.10	2.17 ± 0.09	0.194				
(kg)	*B*	103	265	2.22 ± 0.10	2.17 ± 0.09	0.291				
	C	329	39	2.17 ± 0.09	2.26 ± 0.11	0.268				
CFW	*A*	120	248	1.67 ± 0.08	1.64 ± 0.08	0.464				
(kg)	*B*	103	265	1.66 ± 0.08	1.65 ± 0.08	0.671				
	C	329	39	1.65 ± 0.08	1.65 ± 0.10	0.975				
Yield	*A*	120	248	75.0 ± 1.44	75.7 ± 1.34	0.285				
(%)	*B*	103	265	75.4 ± 1.48	75.6 ± 1.35	0.838				
	C	329	39	75.8 ± 1.34	72.8 ± 1.66	**0.007**				
MFD	*A*	120	248	18.8 ± 0.40	18.6 ± 0.37	0.229				
(µm)	*B*	103	265	18.8 ± 0.41	18.6 ± 0.37	0.399				
	C	329	39	18.5 ± 0.36	20.1 ± 0.45	**<0.001**				
FDSD	*A*	120	248	4.03 ± 0.15	3.93 ± 0.15	0.120	*C*	4.17 ± 0.15	4.16 ± 0.15	0.823
(µm)	*B*	103	265	3.98 ± 0.16	3.94 ± 0.15	0.569				
	C	329	39	3.90 ± 0.14	4.49 ± 0.18	**<0.001**	*A*	3.90 ± 0.14	4.42 ± 0.18	**<0.001**
CVFD	*A*	120	248	21.3 ± 0.49	20.9 ± 0.45	0.127	*C*	21.6 ± 0.50	21.4 ± 0.49	0.428
(%)	*B*	103	265	21.3 ± 0.50	21.0 ± 0.45	0.267				
	C	329	39	20.9 ± 0.45	22.0 ± 0.60	**0.009**	*A*	21.0 ± 0.46	22.0 ± 0.60	**0.023**
MSL	*A*	120	248	83.4 ± 2.82	81.5 ± 2.64	0.116				
(mm)	*B*	103	265	82.0 ± 2.89	81.9 ± 2.65	0.941				
	C	329	39	81.7 ± 2.64	84.0 ± 3.37	0.314				
MSS	*A*	120	248	24.5 ± 2.25	25.3 ± 2.11	0.388				
(N/ktex)	*B*	103	265	25.0 ± 2.30	25.2 ± 2.11	0.858				
	C	329	39	25.3 ± 2.11	23.4 ± 2.69	0.291				
MFC	*A*	120	248	*92.1 ± 3.22*	*89.3 ± 3.01*	*0.053*	*B*	90.1 ± 3.42	88.5 ± 3.05	0.294
(°/mm)	*B*	103	265	86.8 ± 3.28	90.3 ± 3.00	**0.020**	*A*	87.9 ± 3.46	90.7 ± 3.03	0.096
	C	329	39	89.7 ± 3.03	91.9 ± 3.76	0.361				
PF	*A*	120	248	1.96 ± 0.63	1.57 ± 0.60	0.160	*C*	2.62 ± 0.63	2.65 ± 0.62	0.888
(%)	*B*	103	265	1.41 ± 0.64	1.65 ± 0.59	0.423				
	C	329	39	1.41 ± 0.57	3.87 ± 0.71	**<0.001**	*A*	1.38 ± 0.58	3.87 ± 0.71	**<0.001**

[1] Predicted means, standard errors and *P* values from General Linear Models (GLMs). $P \leq 0.05$ are in bold whereas $0.05 < P < 0.1$ are in italics.

[2] GFW – Greasy Fleece Weight; CFW – Clean Fleece Weight; MFD – Mean Fibre Diameter; FDSD – Fibre Diameter Standard Deviation; CVFD – Coefficient of Variation of Fibre Diameter; MSL – Mean Staple Length; MSS – Mean Staple Strength; MFC – Mean Fibre Curvature; PF – Prickle Factor (percentage of fibres over 30 microns).

Variant *B* was associated with increased MFC in the 'single-variant' model, but this was only a trend in the 'multi-variant' model (Table 4-4).

Among the four genotypes (*AF*, *AB*, *BB* and *BC*) that occurred at a frequency of greater than 5%, an effect of genotype on MFD, FDSD, CVFD, PF and MFC was detected. *BC* sheep produced wool of higher MFD, FDSD and PF than *AA*, *AB* and *BB* sheep, and of higher CVFD than *AB* sheep. Wool from *BB* sheep had a higher MFC than wool from *AA* sheep (Table 4-6). Although the overall genotype effect on yield was not significant ($P = 0.056$), a pairwise comparison revealed a different in yield between genotypes *AB* (75.9 ± 1.35%) and *BC* (72.0 ± 1.85%) (Table 4-5).

Table 4-5. The effect of *KRTAP6-1* genotype on various wool traits

Trait[1]	Mean ± SE[2]				*P*
	AA (n=84)	*AB* (n=149)	*BB* (n=96)	*BC* (n=20)	
GFW (kg)	2.23 ± 0.10	2.15 ± 0.09	2.24 ± 0.10	2.26 ± 0.13	0.205
CFW (kg)	1.67 ± 0.08	1.64 ± 0.08	1.69 ± 0.08	1.64 ± 0.10	0.681
Yield (%)	75.4 ± 1.49^{ab}	75.9 ± 1.35^{a}	75.6 ± 1.47^{ab}	72.0 ± 1.85^{b}	*0.056*
MFD (µm)	18.5 ± 0.40^{b}	18.5 ± 0.36^{b}	18.4 ± 0.39^{b}	20.3 ± 0.49^{a}	**<0.001**
FDSD (µm)	3.96 ± 0.16^{b}	3.89 ± 0.14^{b}	3.91 ± 0.15^{b}	4.61 ± 0.19^{a}	**<0.001**
CVFD (%)	21.3 ± 0.51^{ab}	20.8 ± 0.44^{b}	21.3 ± 0.50^{ab}	22.7 ± 0.68^{a}	**0.003**
MSL (mm)	82.8 ± 2.98	81.3 ± 2.69	83.3 ± 2.93	85.8 ± 3.69	0.270
MSS (N/ktex)	24.7 ± 2.32	25.5 ± 2.09	24.7 ± 2.29	22.6 ± 2.88	0.554
MFC (°/mm)	86.9 ± 3.88^{b}	90.7 ± 3.50^{ab}	93.2 ± 3.82^{a}	92.5 ± 4.80^{ab}	**0.028**
PF (%)	1.23 ± 0.63^{b}	1.49 ± 0.57^{b}	1.16 ± 0.62^{b}	4.51 ± 0.78^{a}	**<0.001**

[1] GFW – Greasy Fleece Weight; CFW – Clean Fleece Weight; MFD – Mean Fibre Diameter; FDSD – Fibre Diameter Standard Deviation; CVFD – Coefficient of Variation of Fibre Diameter; MSL – Mean Staple Length; MSS – Mean Staple Strength; MFC – Mean Fibre Curvature; PF – Prickle Factor (percentage of fibres over 30 microns).

[2] Predicted means and standard errors and *P* values from General Linear Models (GLMs). $P \leq 0.05$ are in bold, whereas $0.05 < P < 0.1$ are in italics. Means within rows that do not share a superscript letter are different ($P < 0.05$).

4.4 Discussion

4.4.1 Various Degrees of Correlations between Wool Traits

Overall, the phenotypic correlations observed between the various wool traits were similar to those observed in other studies (Huisman & Brown, 2009; Safari *et al.*, 2007; Wuliji *et al.*, 2001), but with some exceptions. The correlation between MFD and yield was moderate and negative (r = -0.357) in this study, but was negligible in the studies of Wuliji *et al.* (2001) and Safari *et al.* (2007). The negative correlation between CFW and MFC (r = -0.533) found in this study was stronger than that (r = -0.28) reported by Huisman and Brown (2009). In contrast, the correlation between CFW and MFD was negligible (r = -0.074) in this study, compared to the positive estimates reported in other studies (Huisman & Brown, 2009; Safari *et al.*, 2007; Wuliji *et al.*, 2001). The reasons for these discrepancies are not clear, but it may reflect differences in the sheep investigated. In the current study the sheep were one year old Merino × Southdown crosses, with ram lambs being castrated at approximately one month of age, while the sheep investigated in the other studies (Huisman & Brown, 2009; Safari *et al.*, 2007; Wuliji *et al.*, 2001) were mixed age Merinos with no gender information available. In this study, the sheep had a lower CFW (mean 1.73 ± s.d. 0.39 kg) than the Merinos in the other studies (Huisman & Brown, 2009; Safari *et al.*, 2007), and this is not surprising as these sheep were being shorn for the first time. Both age and gender are known to influence wool traits, with differences reported in the correlation of wool traits between young and old sheep, and between male and female sheep (Huisman & Brown, 2009; Lewer *et al.*, 1994; Roldan *et al.*, 2010). As an example, Roldan *et al.* (2010) reported a moderate positive correlation (r = +0.39) between CFW and MFD at first shearing and a negligible correlation (r = -0.04) at second shearing.

4.4.2 Variation in *KRTAP1-2* Mainly Affects Fleece Weight

This study describes *KRTAP1-2* variation in Merino-cross sheep and its association with wool traits. This variation was revealed using PCR-SSCP to screen a small variable region of *KRTAP1-2* and the identification of two new variants of ovine *KRTAP1-2* brings the number of variants identified from nine to eleven. It is expected that more variants may be identified if extended gene regions are analysed and as more sheep from different sires and different breeds are investigated. This suggests that ovine *KRTAP1-2* is a highly polymorphic gene and given that the *KRTAP1* genes are expressed early in fibre development (Powell & Rogers, 1997), the functional significance of this variation on wool traits warrants investigation.

Variation in *KRTAP1-2* was found to be associated with a number of wool traits, including GFW, CFW, Yield, MSL, MSS, FDSD, CVFD, MFC and PF. This suggests that *KRTAP1-2* has an impact on both wool quantity and quality. However, the largest and most enduring effect appeared to be on the wool

quantity traits GFW and CFW, for which a sizeable difference in mean among common genotypes was detected and a difference that reinforced the conclusions drawn from the variant absence/presence models. As an example, the presence of either *A* or *B* in a genotype was associated with an additional 130 or 140 grams of CFW, or about another 8.4 or 9.1% of clean wool. This effect seemed to be associated with the findings for MSL, suggesting the increase in GFW and CFW may have come about because of increased fibre growth, as opposed to an increase in fibre diameter traits. While one of the variants that appeared to be favourable for GFW and CFW (variant *A*) was also found to be associated with a decreased fibre curvature in the single variant model, the effect was small (less than 4 °/mm).

The possibility exists that the effects observed for *KRTAP1-2* may be due to its linkage to other *KRTAPs* or *KRTs* on the same chromosome. Eighteen other *KRTAPs* and *KRTs* have been identified on sheep chromosome 11 near *KRTAP1-2* (Oar_v3.1 reference assembly). These include *KRT27* (NM_001114763), *KRT39* (HQ283083), *KRT40* (HQ283084), *KRTAP3-3* (NM_001009475), *KRTAP3-2* (M21099), *KRTAP1-4* (X01610), *KRTAP1-1* (X01610), *KRTAP1-2*, *KRTAP1-3* (X02925), *KRTAP4-3* (EU239778), *KRTAP4-1* (X73462), *KRT33A* (NM_001199067), *KRT33B* (NM_001199070), *KRT31* (NM_001009445), *KRT38* (NM_001114762), *KRT32* (HQ283078), *KRT35* (NM_001114761), *KRT15* (AJ006277), and *KRT14* (GQ372828), in order from the centromere to the telomere on the long arm of the chromosome. The clustering of so many linked *KRTAPs* and *KRTs* that are all potentially polymorphic and expressed in the wool fibre, makes it difficult to unravel the independent effects of the individual *KRTAPs* and *KRTs*. Nevertheless the results from this study suggest that *KRTAP1-2* alone could potentially act as a gene-marker for wool weights.

The polymorphism found in ovine *KRTAP1-2* suggests that it might be an excellent candidate for a genetic-marker. Although all the *KRTAPs* and *KRTs* that have been investigated are polymorphic, the level of polymorphism seen in *KRTAP1-2* is now higher than any of the other genes, with eleven variants now having been identified compared to between two to nine variants found in other *KRTAPs* and *KRTs* (Gong *et al.*, 2011b; Gong *et al.*, 2014; Gong *et al.*, 2011c; Gong *et al.*, 2012; Itenge-Mweza *et al.*, 2007; Mckenzie *et al.*, 2012; Zhou *et al.*, 2012). The underlying forces driving or sustaining this polymorphism are still poorly understood.

The nature of the polymorphism may also be important regarding the expression of the gene. The majority (seven out of ten) of the SNPs now described in ovine *KRTAP1-2* are either synonymous or located in the 5'-UTR and 3'-UTR of the gene. Synonymous SNPs can affect mRNA stability (Duan *et al.*, 2003) or alter mRNA secondary structure (Nackley *et al.*, 2006). Alternatively, the SNPs could be linked to genetic variation in promoter sequences, transcription factor binding sites or other enhancer elements that potentially affect the expression of the gene.

In wool production systems, both CFW and MFD are economically important traits. A weak positive correlation between these two traits has been reported in other studies (Huisman & Brown, 2009; Safari *et al.*, 2007; Wuliji *et al.*, 2001), but the Trangi QPLU$ project has revealed that selection for either CFW or MFD alone will not necessarily lead to deterioration of the other trait (Mortimer *et al.*, 2006), and that the apparent antagonism described (Huisman & Brown, 2009; Safari *et al.*, 2007; Wuliji *et al.*, 2001) can be overcome and both traits simultaneously improved through genetic selection (Mortimer *et al.*, 2006). In the context of the findings reported here, selection for variant *A* or *B* of *KRTAP1-2* to improve GFW, CFW and yield without affecting MFD, may be possible. This would however require further testing of flocks differing in breed, gender and age.

4.4.3 Variation in *KRTAP6-1* mainly Affect Wool Fibre Diameter

Genetic variation in the ovine KAP6-1 gene and its association with various wool traits in a Merino-cross flock was also described in this study. Variation was not only observed in the nucleotide sequence, but also in the length of coding sequence. This is consistent with the trend of KAP genes being polymorphic. The length variation occurred within a region of repeating nucleotide sequence, and thus notionally, within a region of repetitive polypeptide sequences. This kind of sequence repetition and deletion/insertion that maintains reading frame has also been reported for *KRTAP1-1* and *KRTAP5-4* in sheep (Gong *et al.*, 2010c; Rogers *et al.*, 1994b), and *KRTAP1-n* and *KRTAP4-n* in humans (Kariya *et al.*, 2005; Shimomura *et al.*, 2002). It supports the contention that length variation is a structural hallmark of KAP genes.

The presence/absence GLM models revealed variant *C* was associated with an increase in MFD, FDSD, CVFD and PF, and a decrease in yield. This is consistent with the pairwise comparisons, which revealed that the *C* containing genotype (*BC*) had greater MFD, FDSD and PF than the genotypes (*AA*, *AB* and *BB*), which did not contain *C*, and that *BC* genotypes had a higher CVFD but a lower yield than genotype *AB*. Given that MFD was highly correlated with FDSD and PF, moderately positively correlated with CVFD, and moderately negatively correlated with yield, and that FDSD was highly positively correlated with CVFD (Table 4-1), it is not surprising that associations were observed with all of these traits. The observation of homozygotes of *B* having a higher MFC than homozygotes of *A* is also in agreement with the result from the presence/absence model in which the presence of variant *B* was associated with an increase in MFC.

The finding that variant *C* was associated with increased MFD but not with MFC, and that variant *B* was associated with increased MFC but not with MFD, appears to be unexpected, as it is thought that increasing MFC is associated with finer wool. Although a moderate negative correlation between MFD and MFC has been reported (Taylor *et al.*, 1999), the correlation is dependent on the source and

type of wool (Fish, 2002). Different relationships between MFD and MFC for cashmere of different origins have also been reported (McGregor, 2007). In this study, a moderate positive correlation was found between MFD and MFC (Table 4-1), but no correlation between MFC and MFD was found within Merino wool (GSG Wool Testing Services, n.d.). This suggests that MFC and MFD are independent traits (GSG Wool Testing Services, n.d.), although the precision with which MFC can be measured, and challenges in standardising MFC measurement, suggest caution is needed in interpreting correlations between MFD and MFC.

The deviation from HWE observed in this study is probably due to the non-random mating, and not genotyping error or other factors that may affect variant frequency, although the latter always remains a possibility with any DNA-typing technology. The increase in homozygosity and the decrease in heterozygosity suggests that inbreeding has occurred in this flock. There were only six sires used in this study, with five of them being *AB* genotype and one being *BB* genotype. For this reason, 'sire' was included as a random factor in the GLM models to correct the effect of non-random mating. The high frequency of *B* in the sires would appear to be inconsistent with the frequency of *B* in the ewes, but why this has occurred can only be speculated upon in the absence of detail about selection pressure applied to the rams. It is also conceivable that *B* is just more common in this flock.

Despite variation in *KRTAP6-1* being found to be associated with a number of wool traits, the largest effect appears to be on fibre diameter associated traits including MFD, FDSD, CVFD and PF. This does create a challenge though, because the associations revealed in the models may just reflect the correlation of the various fibre diameter traits, and thus it cannot be ascertained which specific trait, if any, the variation in the gene is actually associated with.

The association between variation in *KRTAP6-1* and wool fibre diameter reported here appears to be in agreement with the finding of Parsons *et al.* (1994b) who reported an association between a *Bam*HI RFLP in a KAP6 gene and wool fibre diameter in a single sire-line. The RFLP described in Parsons *et al.* (1994b) may be linked to the variation reported here for *KRTAP6-1*, but it is difficult to match the *KRTAP6-1* variants with the RFLP variants as a *Bam*HI recognition site cannot be found among the *KRTAP6-1* sequences revealed here, and the location of the *Bam*HI RFLP has not been determined. It could therefore be concluded that this study is the first one to link variation in wool fibre diameter with variation in a specific KAP gene.

While HS- and UHS-KAPs are known to cross-link with IFs via extensive disulphide bonding in the wool fibre (Powell & Rogers, 1997), the function of HGT-KAPs in the fibre matrix remains unknown. Matsunag *et al.* (2013) revealed that HGT-KAPs specifically bind to the head domains of IF proteins

via cation-π interactions, and proposed that they act as a biological factor that regulates the arrangement of IFs. In the orthocortex, IFs are arranged helically with the helical angles progressively increasing from the centre to the periphery of macrofibrils, which is different to the pattern of parallel arrangements in the paracortex and mesocortex (Caldwell *et al.*, 2005). It is possible that KAP6-1 regulates the packing density and helical angles of IFs, and/or the interactions with neighbouring IFs and KAPs, and consequently it affects wool fibre diameter and other wool traits. This is supported by previous studies showing that HGT-KAPs are expressed soon after the synthesis of IF proteins (Rogers, 2004; Rogers *et al.*, 2002), that the content of HGT-KAPs is lower in Lincoln sheep wool than Merino sheep wool (Gillespie, 1990), and that a decrease in wool fibre diameter seen in Merino sheep grown on a restricted diet is associated with an increase in KAP6 proteins (Almeida *et al.*, 2014). It has also been reported that the content of HGT-KAPs is decreased in Merino mutant wool that loses crimp (Li *et al.*, 2009) and that the helical angle of IFs in the orthocortex is associated with fibre curvature (Caldwell *et al.*, 2005).

In the context of variant *C*, the 57 bp deletion would notionally result in the loss of four (20%) aromatic and two (50%) basic amino acid residues (Figure 4-3). Given that aromatic and basic residues are involved in the formation of cation-π interactions (Gallivan & Dougherty, 1999), the reduction of these residues may weaken the binding strength of KAP6-1 and IF proteins. The 57 bp deletion would also lead to the loss of three (33%) cysteine and seven (23%) glycine residues (Figure 4-3). Cysteine is vital for wool growth and is usually the first-limiting amino acid for wool fibre synthesis. It would be expected that variant *C* would favour the synthesis of KAP6-1 protein, but this appears to be in contrast to the findings of Almeida *et al.* (2014) that increasing KAP6 protein levels were associated with finer wool. While the role of glycine in HGT-KAPs has not been established, glycine is the smallest amino acid and lacks a side chain, which may make the HGT-KAPs more flexible and thus better able to form a compact structure with IF proteins. Overall this suggests that the 57mbp deletion seen in variant *C* may affect the structure and/or expression of the KAP6-1 protein, and/or its interaction with IF proteins, and consequently have an effect on wool fibre diameter associated traits. However, it could also be that the association comes about due to linkage with other genes on ovine chromosome 1, not least the other KAP6 genes and/or the KAP7 or KAP8 genes. Further investigation of more sheep from different breeds is required to confirm this finding.

Chapter 5

General Summary and Future Directions

This thesis focused on the identification of KAP genes in sheep and the investigation of variation in both newly identified and some previously identified KAP genes. The sequence information obtained was then used to investigate whether variation in KAP genes was associated with wool traits. To accommodate some emerging issues associated with the diversity of KAP genes, an update nomenclature for KAPs and *KRTAPs* was introduced.

Seven new ovine KAP genes (*KRTAP1-2, KRTAP6-2, KRTAP6-3, KRTAP8-2, KRTAP11-1, KRTAP13-3* and *KRTAP24-1*) were identified (Chapter 2), increasing the number of KAP genes identified in sheep from 16 to 23. Despite this, the ovine KAP genes identified so far probably represent only a small proportion of the KAP genes in the sheep genome, given that there are 88 KAP genes in humans (Rogers & Schweizer, 2005). The identification of a second member of the KAP8 gene family in sheep, a family member which is absent in humans (Chapter 2), suggests that sheep may have either more or different KAP genes than humans. Accordingly further study is required to characterise more KAP genes in the sheep genome, including those that are human homologues and those that are present in sheep, but absent in humans and other species. Recent progress in sheep genome sequencing and the recent identification of KAP genes in other species, for example, goat KAP9.2 (Wang *et al.*, 2012) and human KAP2 (Fujikawa *et al.*, 2012), will facilitate this further study.

All the KAP genes investigated in this study were polymorphic, but the level and nature of polymorphism varied considerably. The number of allelic variants identified ranged from two alleles for *KRTAP7-1* and *KRTAP8-2* to eleven alleles for *KRTAP1-2*; and the variation detected included SNPs and length variation (Chapter 2). In addition, non-synonymous SNPs were predominated over synonymous SNPs in *KRTAP1-4, KRTAP5-4, KRTAP13-3* and *KRTAP24-1*, whereas the opposite was observed in *KRTAP1-2*, *KRTAP8-1* and *KRTAP11-1* (Chapter 2). The mechanisms driving the accumulation of variation are currently unknown. Investigation of the upstream and downstream regions of the KAP genes, may shed some light on the evolution of KAP gene variation.

The emerging diversity in KAP genes and the impact of species of origin prompted me to revise the current nomenclature for mammalian KAPs/KRTAPs. It is proposed to use sp-KAPm-nL*x for the protein and sp-*KRTAPm-n(p/L)*x* for the gene, to replace the current use of KAPm.nxpL for the protein and *KRTAPm.nxpL* for the gene (Chapter 3). The revised nomenclature was not a major change from the current nomenclature, but included a species identifier to constitute an informative

KAP naming system, and had a minor change in the use and order of “p”, “L” and “x”, which should have minimal impact on current publications and databases.

Variation in *KRTAP1-2* was found to be associated with fleece weight, while variation in *KRTAP6-1* was associated with wool fibre diameter (Chapter 4). These findings could form the basis for the development of gene-markers for improving fleece weight and fibre diameter. However, before such gene-markers can be developed, these associations need to be confirmed in larger flocks of sheep and in sheep from different breeds. It would be interesting to investigate whether other KAP genes near *KRTAP1-2* or *KRTAP6-1* have a similar effect on the traits. Should such an effect be identified, extended haplotype analysis of these KAP gene chromosomal regions should be considered and association studies at the haplotype level may be a better option for finding genetic markers for these wool traits. Future studies should also be carried out to investigate whether variation detected in KAP genes affects gene expression or protein structure. These studies should also enable a better understanding of how KAP genes affect wool traits.

The different effects of *KRTAP1-2* and *KRTAP6-1* on fleece weight and fibre diameter observed in this study (Chapter 4), suggests that the two traits can be independently selected for and will not detract from each other. This is in contrast to previous findings (Huisman and Brown, 2009; Safari *et al.*, 2007; Wuliji *et al.*, 2001), but is consistent with the finding of the Trangi QPLU$ project which revealed that clean fleece weight or mean fibre diameter can be simultaneously improved by genetic selection (Mortimer *et al.*, 2006). Given that *KRTAP1-2* and *KRTAP6-1* are located on different chromosomes, selective breeding for these traits using gene markers based on variation in thee loci can be achieved should the effects of these KAP genes be confirmed. However, before such gene markers can be developed, the associations identified in this study need to be confirmed in a large number of sheep and in sheep from other breeds, and then tracked from first shear sheep to older sheep.

References

Almeida, A. M., Plowman, J. E., Harland, D. P., Thomas, A., Kilminster, T., Scanlon, T., Milton, J., Greeff, J., Oldham, C., and Clerens, S. (2014). Influence of feed restriction on the wool proteome: A combined iTRAQ and fiber structural study. *Journal of Proteomics, 103*, 170-177.

Angel, C., Beare, S., and Zwart, A. (1990). Product characteristics and arbitrage in the Australian and New Zealand wool markets. *Australian Journal of Agricultural Economics, 34*(1), 67-79.

Aylan-Parker, J., and McGregor, B. A. (2002). Optimising sampling techniques and estimating sampling variance of fleece quality attributes in alpacas. *Small Ruminant Research, 44*(1), 53-64.

AWI. (n.d.). Browse data. Retrieved from http://www.wool.com/

Beh, K. J., Callaghan, M. J., Leish, Z., Hulme, D. J., Lenane, I., and Maddox, J. F. (2001). A genome scan for QTL affecting fleece and wool traits in Merino sheep. *Wool Technolgoy and Sheep Breeding, 49*, 88-97.

Benavides, M., and Maher, A. (2000). Quantitative genetic studies on wool yellowing in Corriedale sheep. II. Clean wool colour and wool production traits: genetic parameter estimates and economic returns. *Crop and Pasture Science, 51*(2), 191-196.

Botha, A. F., and Hunter, L. (2010). The measurement of wool fibre properties and their effect on worsted processing performance and product quality. Part 1: The objective measurement of wool fibre properties. *Textile Progress, 42*(4), 227-339.

Bowden, P. E., Stark, H. J., Breitkreutz, D., and Fusenig, N. E. (1987). Expression and modification of keratins during terminal differentiation of mammalian epidermis. *Current Topics in Developmental Biology, 22*, 35-68.

Bowman, J. C. (Ed.) (1974). *An introduction to Animal Breeding.* London: Edward Arnold.

Brown, D., Crook, B., and Purvis, I. (2002). Differences in fibre diameter profile characteristics in wool staples from Merino sheep and their relationship with staple strength between years, environments, and bloodlines. *Crop and Pasture Science, 53*(4), 481-491.

Brown, G. H., Turner, H. N., Young, S. S. Y., and Dolling, C. H. S. (1966). Vital statistics for an experimental flock of Merino sheep. III. Factors affecting wool and body characteristics,

including the effect of age of ewe and its possible interaction with method of selection. *Crop and Pasture Science, 17*(4), 557-581.

Bush, W. S., and Moore, J. H. (2012). Genome-wide association studies. *PLoS Computational Biology, 8*(12), e1002822.

Byun, S. O., Fang, Q., Zhou, H., and Hickford, J. G. H. (2009). An effective method for silver-staining DNA in large numbers of polyacrylamide gels. *Analytical Biochemistry, 385*(1), 174-175.

Caldwell, J. P., Mastronarde, D. N., Woods, J. L., and Bryson, W. G. (2005). The three-dimensional arrangement of intermediate filaments in Romney wool cortical cells. *Journal of Structural Biology, 151*(3), 298-305.

Campbell, M. E., Whiteley, K. J., and Gillespie, J. M. (1972). Compositional studies of high-and low-crimp wools. *Australian Journal of Biological Sciences, 25*(5), 977-988.

Cole, S. E., and Reeves, R. H. (1998). A cluster of keratin-associated proteins on mouse chromosome 10 in the region of conserved linkage with human chromosome 21. *Genomics, 54*(3), 437-442.

Conner, B. J., Reyes, A. A., Morin, C., Itakura, K., Teplitz, R., and Wallace, R. B. (1983). Detection of sickle cell beta S-globin allele by hybridization with synthetic oligonucleotides. *Proceedings of the National Academy of Sciences, 80*(1), 278-282.

Cottle, D. J. (Ed.) (1991). *Australian Sheep and Wool Handbook.* Melbourne: Inkata Press.

Cottle, D. J. (Ed.) (2010). *International Sheep and Wool Handbook.* Nottingham, UK: Nottingham University Press.

Crawford, A. M., Dodds, K. G., Ede, A. J., Pierson, C. A., Montgomery, G. W., Garmonsway, H. G., Beattie, A. E., Davies, K., Maddox, J. F., and Kappes, S. W. (1995). An autosomal genetic linkage map of the sheep genome. *Genetics, 140*(2), 703-724.

Crewther, W., Dowling, L., Gough, K., Inglis, A., McKern, N., Sparrow, L., and Woods, E. (1975*).* The low-sulphur proteins of wool: studies on their classification, characterization, primary and secondary structure*. Proeceedings of the 5th International Wool Textile Research Conference*. *2*, 233-242.

Crewther, W. G., Inglis, A. S., and McKern, N. M. (1978). Amino acid sequences of alpha-helical segments from S-carboxymethylkerateine-A. Complete sequence of a type-II segment. *Biochemical Journal, 173(2)*, 365-371.

David, V., and Deutch, A. (1992). Detection of bovine αs1-casein genomic variants using the allele-specific polymerase chain reaction. *Animal Genetics, 23*(5), 425-429.

den Dunnen, J. T., and Antonarakis, S. E. (2000). Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Human Mutation*, *15*(1), 7-12.

Dopheide, T. A. (1973). The primary structure of a protein, component 0.62, rich in glycine and aromatic residues, obtained from wool keratin. *European Journal of Biochemistry, 34*(1), 120-124.

Dowling, L. M., Crewther, W. G., and Inglis, A. S. (1986). The primary structure of component 8c-1, a subunit protein of intermediate filaments in wool keratin. Relationships with proteins from other intermediate filaments. *Biochemical Journal, 236(3)*, 695-703.

Duan, J., Wainwright, M. S., Comeron, J. M., Saitou, N., Sanders, A. R., Gelernter, J., and Gejman, P. V. (2003). Synonymous mutations in the human dopamine receptor D2 (DRD2) affect mRNA stability and synthesis of the receptor. *Human Molecular Genetics, 12*(3), 205-216.

Elleman, T. C. (1971). Amino-acid sequence of a high-sulphur protein from wool. *Nature, 234*(48), 148-148.

Elleman, T. C. (1972a). The amino acid sequence of protein SCMK-B2C from the high-sulphur fraction of wool keratin. *Biochemical Journal, 128(5)*, 1229-1239.

Elleman, T. C. (1972b). The amino acid sequence of protein SCMK-B2A from the high-sulphur fraction of wool keratin. *Biochemical Journal, 130(3)*, 833-845.

Elleman, T. C., and Dopheide, T. A. (1972). The sequence of SCMK-B2B, a high-sulfur protein from wool keratin. *Journal of Biological Chemistry, 247*(12), 3900-3909.

Engelke, D. R., Hoener, P. A., and Collins, F. S. (1988). Direct sequencing of enzymatically amplified human genomic DNA. *Proceedings of the National Academy of Sciences, 85*(2), 544-548.

FAOSTAT. (2015). Browse data. Retrieved from http://faostat3.fao.org

Fish, V. (2002). Measuring fibre curvature: Key issues. *Wool Technology and Sheep Breeding, 50*(4), 792-799.

Fratini, A., Powell, B. C., and Rogers, G. E. (1993). Sequence, expression, and evolutionary conservation of a gene encoding a glycine tyrosine-rich keratin-associated protein of hair. *Journal of Biological Chemistry, 268*(6), 4511-4518.

Fratini, A., Powell, B. C., Hynd, P. I., Keough, R. A., and Rogers, G. E. (1994). Dietary cysteine regulates the levels of messenger-rnas encoding a family of cysteine-rich proteins of wool. *Journal of Investigative Dermatology, 102*(2), 178-185.

Frenkel, M. J., Powell, B. C., Ward, K. A., Sleigh, M. J., and Rogers, G. E. (1989). The keratin BIIIB gene family: Isolation of cDNA clones and structure of a gene and a related pseudogene. *Genomics, 4*(2), 182-191.

Fuchs, E., and Weber, K. (1994). Intermediate filaments: structure, dynamics, function and disease. *Annual Review of Biochemistry, 63*(1), 345-382.

Fujikawa, H., Fujimoto, A., Farooq, M., Ito, M., and Shimomura, Y. (2012). Characterization of the human hair keratin-associated protein 2 (KRTAP2) gene family. *Journal of Investigative Dermatology, 132*(7), 1806-1813.

Gallivan, J. P., and Dougherty, D. A. (1999). Cation-π interactions in structural biology. *Proceedings of the National Academy of Sciences, 96*(17), 9459-9464.

Gifford, D., Ponzoni, R., Ancell, P., Hynd, P., Walkley, J., and Grimson, R. (1995). Genetic studies on wool quality and skin characters of the Merino. *Wool Technology and Sheep Breeding, 43*(1), 24-29.

Gillespie, J. (1963). The isolation and properties of some soluble proteins from wool. V. The isolation of the high-sulphur protein SCMKB2. *Australian Journal of Biological Sciences, 16*(1), 241-251.

Gillespie, J. (1966). The dietary-regulated biosynthesis of high-sulphur wool proteins. *Biochemical Journal, 98(3)*, 669-677.

Gillespie, J., and Broad, A. (1972). Ultra-high-sulphur proteins in the hairs of the artiodactyla. *Australian Journal of Biological Sciences, 25*(1), 139-146.

Gillespie, J. M. (1983). The structural proteins of hair: isolation, characterization and regulation of biosynthesis. In LA. Goldsmith (Ed.), *Biochemistry and Physiology of the Skin*, volume 1 (pp.475-510): Oxford University Press.

Gillespie, J. M. (1990). The proteins of hair and other hard α-keratins. In R. D. Goldman, P. M. Steinert (Ed.), *Cellular and Molecular Biology of Intermediate Filaments* (pp. 95-128). New York: Plenum Press.

Gillespie, J. M., and Darskus, R. L. (1971). Relation between the tyrosine content of various wools and their content of a class of proteins rich in tyrosine and glycine. *Australian Journal of Biological Sciences, 24*(4), 1189-1198.

Goddard, D. R., and Michaelis, L. (1935). Derivatives of keratin. *Journal of Biological Chemistry, 112*(1), 361-371.

Gong, H., Zhou, H., and Hickford, J. G. (2010a). Polymorphism of the ovine keratin-associated protein 1-4 gene (KRTAP1-4). *Molecular Biology Reports, 37*(7), 3377-3380.

Gong, H., Zhou, H., and Hickford, J. G. H. (2010b). Polymorphism of the ovine keratin-associated protein 1-4 gene (*KRTAP1-4*). *Molecular Biology Reports, 37*(7), 3377-3380.

Gong, H., Zhou, H., and Hickford, J. G. H. (2011a). Diversity of the glycine/tyrosine-rich keratin-associated protein 6 gene (KAP6) family in sheep. *Molecular Biology Reports, 38*(1), 31-35.

Gong, H., Zhou, H., Dyer, J. M., and Hickford, J. G. (2011b). Identification of the ovine KAP11-1 gene (*KRTAP11-1*) and genetic variation in its coding sequence. *Molecular Biology Reports, 38*(8), 5429-5433.

Gong, H., Zhou, H., Dyer, J. M., and Hickford, J. G. (2014). The sheep KAP8-2 gene, a new KAP8 family member that is absent in humans. *SpringerPlus, 3*(1), 1-5.

Gong, H., Zhou, H., Plowman, J. E., Dyer, J. M., and Hickford, J. G. (2010c). Analysis of variation in the ovine ultra-high sulphur keratin-associated protein KAP5-4 gene using PCR-SSCP technique. *Electrophoresis, 31*(21), 3545-3547.

Gong, H., Zhou, H., Dyer, J. M., Plowman, J. E., and Hickford, J. G. H. (2011c). Identification of the keratin-associated protein 13-3 (KAP13-3) gene in sheep. *Open Journal of Genetics, 1*(3), 60-64.

Gong, H., Zhou, H., Plowman, J. E., Dyer, J. M., and Hickford, J. G. H. (2011d). Search for variation in the ovine KAP7-1 and KAP8-1 genes using polymerase chain reaction-single-stranded conformational polymorphism screening. *DNA and Cell Biology, 31*(3), 367-370.

Gong, H., Zhou, H., Yu, Z., Dyer, J., Plowman, J. E., and Hickford, J. (2011e). Identification of the ovine keratin-associated protein KAP1-2 gene (*KRTAP1-2*). *Experimental Dermatology, 20*(10), 815-819.

Gong, H., Zhou, H., McKenzie, G. W., Yu, Z., Clerens, S., Dyer, J. M., Plowman, J. E., Wright, M. W., Arora, R., Bawden, C. S., Chen, Y., Li, J., and Hickford, J. G. H. (2012). An Updated Nomenclature for Keratin-Associated Proteins (KAPs). *International Journal of Biological Sciences, 8*(2), 258-264.

Gough, K. H., Inglis, A. S., and Crewther, W. G. (1978). Amino acid sequences of alpha-helical segments from S-carbosymethylkerateine-A. Complete sequence of a type-I segment. *Biochemical Journal, 173(2)*, 373-385.

GSG Wool Testing Services. (n.d.). Fibre curvature. Retrieved from http://www.sgs.co.nz/~/media/Local/New%20Zealand/Documents/Technical%20Documents/Technical%20Bulletins/Wool%20Testing%20Info%20Bulletins/SGS-AGRI-5-5-Fibre-Curvature-A4-EN-11-V1.pdf

Gygi, S. P., Rochon, Y., Franza, B. R., and Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. *Molecular and Cellular Biology, 19*(3), 1720-1730.

Hardy, M. H., and Lyne, A. G. (1956). The pre-natal development of wool follicles in Merino sheep. *Australian Journal of Biological Sciences, 9*(3), 423-441.

Harrap, B. (1963). The conformation of a soluble wool keratin derivative. *Australian Journal of Biological Sciences, 16*(1), 231-240.

Hart, G. W., and West, C. M. (2009). Nucleocytoplasmic glycosylation. In A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart & M. E. Etzler (Eds.), *Essentials of Glycobiology* (2nd edition ed., pp. 784pp). New York: Cold Spring Harbor Laboraotry Press.

Hawker, H., Dodds, K., Andrews, R., and McEwan, J. (1988). Production and characteristics of wool from the hogget progeny of sheep intensively screened for fleece weight. *Proceedings of the New Zealand Society of Animal Production*, 48, 207-212.

Hayashi, K. (1991). PCR-SSCP: a simple and sensitive method for detection of mutations in the genomic DNA. *PCR Methods and Applications,* 1, 34-38.

Hayashi, K. (1992). PCR-SSCP: a method for detection of mutations. *Genetic Analysis: Biomolecular Engineering, 9*(3), 73-79.

Haylett, T., and Swart, L. (1969). Studies on the High-Sulfur Proteins of Reduced Merino Wool Part III: The Amino-Acid Sequence of Protein SCMKB-IIIB2. *Textile Research Journal, 39*(10), 917-929.

Haylett, T., Swart, L., and Parris, D. (1971). Studies on the high-sulphur proteins of reduced Merino wool. Amino acid sequence of protein SCMBK-IIIB3. *Biochemical Journal, 123(2)*, 191-200.

Henderson, A. E. (1968). *Growing Better Wool.* Wellington, New Zealand: A. H. and A. W. Reed.

Hesse, M., Magin, T. M., and Weber, K. (2001). Genes for intermediate filament proteins and the draft sequence of the human genome novel keratin genes and a surprisingly high number of pseudogenes related to keratin genes 8 and 18. *Journal of Cell Science, 114*(14), 2569-2575.

Hesse, M., Zimek, A., Weber, K., and Magin, T. M. (2004). Comprehensive analysis of keratin gene clusters in humans and rodents. *European Journal of Cell Biology, 83*(1), 19-26.

Hirschhorn, J. N., and Daly, M. J. (2005). Genome-wide association studies for common diseases and complex traits. *Nature Reviews Genetics, 6*(2), 95-108.

Hocker, H. (2002). Fibre morphology. In W. S.Simpson & G. H. Crwashaw (Eds.), *Wool: Science and Technology* (pp. 60-79). Cambridge, England: Woodhead Publishing Ltd.

Horio, M., and Kondo, T. (1953). Crimping of wool fibers. *Textile Research Journal, 23*(6), 373-386.

Huh, N.-H., Kashiwagi, M., Konishi, C., Hashimoto, Y., Kohno, Y., Nomura, S., and Kuroki, T. (1994). Isolation and characterization of a novel hair follicle-specific gene, Hacl-1. *Journal of Investigative Dermatology, 102*(5), 716-720.

Huisman, A., and Brown, D. (2009). Genetic parameters for bodyweight, wool, and disease resistance and reproduction traits in Merino sheep. 3. Genetic relationships between ultrasound scan traits and other traits. *Animal Production Science, 49*(4), 283-288.

Itenge-Mweza, T. O. (2007). *Identification of genetic markers associated with wool quality traits in Merino sheep.* PhD thesis, Lincoln University, New Zealand.

Itenge-Mweza, T. O., Forrest, R. H., McKenzie, G. W., Hogan, A., Abbott, J., Amoafo, O., and Hickford, J. G. (2007). Polymorphism of the KAP1.1, KAP1.3 and K33 genes in Merino sheep. *Molecular and Cellular Probes, 21*(5-6), 338-342.

Itenge, T., Hickford, J., Forrest, R., McKenzie, G., and Frampton, C. (2010). Association of variation in the ovine KAP1. 1, KAP1. 3 and K33 genes with wool traits. *International Journal of Sheep and Wool Science, 58*(1), 1-20.

IWTO. (2015). Browse data. Retrived from http://www.iwto.org/

IWTO. (2013). IWTO-28: Determination by the Airflow Method of the Mean Fibre Diameter of Core Samples of Raw Wool.

Jenkins, B. J., and Powell, B. C. (1994). Differential expression of genes encoding a cysteine-rich keratin family in the hair cuticle. *Journal of Investigative Dermatology, 103*(3), 310-317.

Jin, M., Wang, L., Li, S., Xing, M. X., and Zhang, X. (2011). Characterization and expression analysis of KAP7.1, KAP8.2 gene in Liaoning new-breeding cashmere goat hair follicle. *Molecular Biology Reports, 38*(5), 3023-3028.

Johansson, I., and Rendel, J. (1968). *Genetics and Animal Breeding*. Edingburgh: Oliver and Boyd.

Kariya, N., Shimomura, Y., and Ito, M. (2005). Size polymorphisms in the human ultrahigh sulfur hair keratin-associated protein 4, KAP4, gene family. *Journal of Investigative Dermatology, 124*(6), 1111-1118.

Ku, N. O., Liao, J., Chou, C. F., and Omary, M. B. (1996). Implications of intermediate filament protein phosphorylation. *Cancer and Metastasis Reviews, 15*(4), 429-444.

Kuczek, E., and Rogers, G. E. (1985). Sheep keratins: characterization of cDNA clones for the glycine+ tyrosine-rich wool proteins using a synthetic probe. *European journal of biochemistry, 146*(1), 89-93.

Kuczek, E. S., and Rogers, G. E. (1987). Sheep wool (glycine+ tyrosine)-rich keratin genes. *European Journal of Biochemistry, 166*(1), 79-85.

Kuhn, F., Lassing, C., Range, A., Mueller, M., Hunziker, T., Ziemiecki, A., and Andres, A. C. (1999). Pmg-1 and pmg-2 constitute a novel family of KAP genes differentially expressed during skin and mammary gland development. *Mechanisms of Development, 86*(1), 193-196.

Kwon, J. M., and Goate, A. M. (2000). The candidate gene approach. *Alcohol Research and Health, 24*(3), 164-168.

Langbein, L., Rogers, M. A., Winter, H., Praetzel, S., and Schweizer, J. (2001). The catalog of human hair keratins II. Expression of the six type II members in the hair follicle and the combined catalog of human type I and II keratins. *Journal of Biological Chemistry, 276*(37), 35123-35132.

Langbein, L., Rogers, M. A., Praetzel-Wunder, S., Böckler, D., Schirmacher, P., and Schweizer, J. (2007). Novel type I hair keratins K39 and K40 are the last to be expressed in differentiation of the hair: completion of the human hair keratin catalog. *Journal of Investigative Dermatology, 127*(6), 1532-1535.

Langbein, L., Rogers, M. A., Winter, H., Praetzel, S., Beckhaus, U., Rackwitz, H. R., and Schweizer, J. (1999). The catalog of human hair keratins I. Expression of the nine type I members in the hair follicle. *Journal of Biological Chemistry, 274*(28), 19874-19884.

Leeder, J. D. (1984). *Wool: nature's wonder fibre*. Victoria, Australia: Australasian Textiles Publishers.

Leng, R. A., and Stephenson, S. K. (1965). Glucose and acetate metabolism by isolated sheep wool follicles. *Archives of Biochemistry and Biophysics, 110*(1), 8-15.

Lewer, R., Woolaston, R., and Howe, R. (1994). Studies on Western Australian Merino sheep. II. Genetic and phenotypic parameter estimates for objectively measured traits on ram and ewe hoggets using different model types. *Crop and Pasture Science, 45*(4), 829-840.

Li, S. W., Ouyang, H. S., Rogers, G. E., and Bawden, C. S. (2009). Characterization of the structural and molecular defects in fibres and follicles of the merino felting lustre mutant. *Experimental Dermatology, 18*(2), 134-142.

Lindley, H., and Elleman, T. C. (1972). The preparation and properties of a group of proteins from the high-sulphur fraction of wool. *Biochemical Journal, 128(4)*, 859-867.

Liu, W., Fang, Y., Fang, G., Wang, M., Yu, H., Li, X., Feng, T., Chen, H., and Yu, S. (2009). Missense mutation of KAP16. 6 gene on three goat breeds in China and their associations with Cashmere production traits. *Journal of Animal and Veterinary Advances, 8(12)*, 2713-2718.

Livak, K. J. (1999). Allelic discrimination using fluorogenic probes and the 5′ nuclease assay. *Genetic Analysis: Biomolecular Engineering, 14*(5), 143-149.

Livestock Genomics. (2011). Data retrived from http://www.livestockgenomics.csiro.au/sheep/oar2.0.php

MacKinnon, P., Powell, B., and Rogers, G. (1990). Structure and expression of genes for a class of cysteine-rich proteins of the cuticle layers of differentiating wool and hair follicles. *The Journal of Cell Biology, 111*(6), 2587-2600.

Malau-Aduli, A. E. O., and Deng Akuoch, D. J. (2010). Wool comfort factor variation in Australian crossbred sheep. *Journal of Animal Science, 88*(E-Supplement 2), 860 Abstr.

Marshall, R. C. (1981). Analysis of the proteins from single wool fibers by two-dimensional polyacrylamide gel electrophoresis. *Textile Research Journal, 51*(2), 106-108.

Marshall, R. C., Orwin, D. F., and Gillespie, J. M. (1991). Structure and biochemistry of mammalian hard keratin. *Electron Microscopy Reviews, 4*(1), 47-83.

Matsunaga, R., Abe, R., Ishii, D., Watanabe, S. I., Kiyoshi, M., Nöcker, B., Tsuchiya, M., and Tsumoto, K. (2013). Bidirectional binding property of high glycine-tyrosine keratin-associated protein contributes to the mechanical strength and shape of hair. *Journal of Structural Biology, 183*(3), 484-494.

McGregor, B. (2007). Cashmere fibre crimp, crimp form and fibre curvature. *International Journal of Sheep and Wool Science, 55*(1), 106-129.

Mckenzie, G. W., Arora, R., and Hickford, J. G. (2012). Genetic variation in the 5′ UTR of the KRT2. 13 gene of sheep. *Animal Science Journal, 83*(3), 194-198.

McLaren, R. J., Rogers, G. R., Davies, K. P., Maddox, J. F., and Montgomery, G. W. (1997). Linkage mapping of wool keratin and keratin-associated protein genes in sheep. *Mammalian Genome, 8*(12), 938-940.

Miklos, G. L. G., and Maleszka, R. (2001). Protein functions and biological contexts. *Proteomics, 1*(2), 169-178.

Miterski, B., Kruger, R., Wintermeyer, P., and Epplen, J. T. (2000). PCR/SSCP detects reliably and efficiently DNA sequence variations in large scale screening projects. *Combinatorial Chemistry & High Throughput Screening, 3*(3), 211-218.

Mishra, S. P. (2000). *A Text Book of Fibre Science and Technology*. New Delhi: New Age International (P) Ltd.

Moll, R., Divo, M., and Langbein, L. (2008). The human keratins: biology and pathology. *Histochemistry and Cell Biology, 129*(6), 705-733.

Mortimer, S., Taylor, P., Atkins, K., and Pope, C. (2006). The Trangie QPLU$ selection lines: responses in clean fleece weight and fibre diameter on completion of ten rounds of selection. In C. E.

Pope (Ed.), *Proceedings of the Trangie QPLUS Open Day (*pp. 7-11). Trangie Agricultural Research Centre, Trangie, Australia, 11 May 2006.

Mortimer, S. I., and Atkins, K. (1989). Genetic evaluation of production traits between and within flocks of Merino sheep. I. Hogget fleece weights, body weight and wool quality. *Crop and Pasture Science, 40*(2), 433-443.

Myers, R. M., Fischer, S. G., Lerman, L. S., and Maniatis, T. (1985). Nearly all single base substitutions in DNA fragments joined to a GC-clamp can be detected by denaturing gradient gel electrophoresis. *Nucleic Acids Research, 13*(9), 3131-3145.

Nackley, A., Shabalina, S., Tchivileva, I., Satterfield, K., Korchynskyi, O., Makarov, S., Maixner, W., and Diatchenko, L. (2006). Human catechol-O-methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure. *Science, 314*(5807), 1930-1933.

Nataraj, A. J., Olivos-Glander, I., Kusukawa, N., and Highsmith Jr, W. E. (1999). Single-strand conformation polymorphism and heteroduplex analysis for gel-based mutation detection. *Electrophoresis, 20*(6), 1177-1185.

Naylor, G., Phillips, D., and Veitch, C. (1995). The relative importance of mean diameter and coefficient of variation of sale lots in determining the potential skin comfort of wool fabrics. *Wool Technology and Sheep Breeding, 43(1),* 69-82,

Naylor, G. R. S. (2010). Fabric-evoked prickle in worsted spun single jersey fabrics. Part 4: Extension from wool to Optim™fine fiber. *Textile Research Journal, 80,* 537-547.

Nei, M. (2005). Selectionism and neutralism in molecular evolution. *Molecular Biology and Evolution, 22*(12), 2318-2342.

Orita, M., Suzuki, Y., Sekiya, T., and Hayashi, K. (1989). Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics, 5*(4), 874-879.

Orwin, D. F. G. (1989). Variations in wool follicle morphology. In G. E. Rodgers, P. J. Reis, K. A. Ward & R. C. Marshall (Eds.), *The Biology of Wool and Hair* (pp. 227-242). London: Chapman & Hall.

Parsons, Y. M., Piper, L. R., and Cooper, D. W. (1994a). Linkage relationships between keratin-associated protein (KRTAP) genes and growth hormone in sheep. *Genomics, 20*(3), 500-502.

Parsons, Y. M., Cooper, D. W., and Piper, L. R. (1994b). Evidence of linkage between high-glycine-tyrosine keratin gene loci and wool fiber diameter in a merino half-sib family. *Animal Genetics, 25*(2), 105-108.

Ratner, B. (n.d.) The Correlation Coefficient: Definition. Retrived from http://www.dmstat1.com/res/TheCorrelationCoefficientDefined.html

Philpott, M. P., and Kealey, T. (1991). Metabolic studies on isolated hair follicles: hair follicles engaged in aerobic glycolysis and do not demonstrate the glucose fatty acid cycle. *Journal of Investigative Dermatology, 96*(6), 875-879.

Pickering, N., Blair, H., Hickson, R., Dodds, K., Johnson, P., and McEwan, J. (2013). Genetic relationships between dagginess, breech bareness, and wool traits in New Zealand dual-purpose sheep. *Journal of Animal Science, 91*(10), 4578-4588.

Plowman, J. E. (2003). Proteomic database of wool components. *Journal of Chromatography B, 787*(1), 63-76.

Plowman, J. E., Deb-Choudhury, S., Clerens, S., Thomas, A., Cornellison, C. D., and Dyer, J. M. (2012). Unravelling the proteome of wool: Towards markers of wool quality traits. *Journal of Proteomics, 75*(14), 4315-4324.

Popescu, C., and Höcker, H. (2007). Hair - the most sophisticated biological composite material. *Chemical Society Reviews, 36*(8), 1282-1291.

Powell, B., and Rogers, G. (1994). Differentiation in hard keratin tissues: hair and related structures. In I. E. Leigh & E. B. Lane (Eds.), *The Keratinocyte Handbook* (pp. 401-436). Cambridge, UK: Cambridge University Press.

Powell, B., Crocker, L., and Rogers, G. (1992). Hair follicle differentiation: expression, structure and evolutionary conservation of the hair type II keratin intermediate filament gene family. *Development, 114*(2), 417-433.

Powell, B. C., and Rogers, G. E. (1986). Hair keratin: composition, structure and biogenesis. In J. Bereiter-Hahn, A. G. Matoltsy & K. S. Richards (Eds.), *Biology of the Integument* (pp. 696-721). Berlin: Springer-Verlag.

Powell, B. C., and Rogers, G. E. (1990). Hard keratin IF and associated proteins. In R. D. Goldman & P. M. Steinert (Eds.), *Cellular and Molecular Biology of Intermediate Filaments* (pp. 267-300). New York: Plenum.

Powell, B. C., and Rogers, G. E. (1997). The role of keratin proteins and their genes in the growth, structure and properties of hair. In P. Jollès, H. Zahn & H. Höcker (Series Eds.), *Formation and Structure of Human Hair* (pp. 59-148). Boston, MA: Birkhäuser Verlag.

Powell, B. C., Crocker, L. A., and Rogers, G. E. (1993). Complete sequence of a hair-like intermediate filament type II keratin gene. *DNA Sequence, 3*(6), 401-405.

Powell, B. C., Arthur, J., and Nesci, A. (1995). Characterization of a gene encoding a cysteine-rich keratin associated protein synthesized late in rabbit hair follicle differentiation. *Differentiation, 58*(3), 227-232.

Powell, B. C., Sleigh, M. J., Ward, K. A., and Rogers, G. E. (1983). Mammalian keratin gene families: organisation of genes coding for the B2 high-sulphur proteins of sheep wool. *Nucleic Acids Research, 11*(16), 5327-5346.

Reid, T., and Botica, K. (1995). Relationship between wool colour and fibre diameter in four breeds of sheep. *Proceedings of the New Zealand Society of Animal Production, 55,* 50–53.

Reis, P. J. (1979). Effects of amino acids on the growth and properties of wool. In J. L. Black & P. J. Reis (Eds.), *Physiological and Environmental Limitations to Wool Growth* (pp. 223–242). Armidale, New South Wales: University of New England Publishing Unit.

Rogers, G., Hickford, J., and Bickerstaffe, R. (1994a). A potential QTL for wool strength located on ovine chromosome 11. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production* (pp. 291-294). 7-12 August; Guelph.

Rogers, G. E. (1959). Electron microscope studies of hair and wool. *Annals of the New York Academy of Sciences, 83*(3), 378-399.

Rogers, G. E. (2004). Hair follicle differentiation and regulation. *International Journal of Developmental Biology, 48*(2-3), 163-170.

Rogers, G. E. (2006). Biology of the wool follicle: an excursion into a unique tissue interaction system waiting to be re-discovered. *Experimental Dermatology, 15*(12), 931-949.

Rogers, G. E., and Powell, B. C. (1993). Organization and expression of hair follicle genes. *Journal of Investigative Dermatology, 101*, 50S-55S.

Rogers, G. R., Hickford, J. G. H., and Bickerstaffe, R. (1994b). Polymorphism in two genes for B2 high sulfur proteins of wool. *Animal Genetics, 25*(6), 407-415.

Rogers, M. A., and Schweizer, J. (2005). Human KAP genes, only the half of it? Extensive size polymorphisms in hair keratin-associated protein genes. *Journal of Investigative Dermatology, 124*(6), vii-ix.

Rogers, M. A., Langbein, L., Praetzel Wunder, S., and Giehl, K. (2008). Characterization and expression analysis of the hair keratin associated protein KAP26. 1. *British Journal of Dermatology, 159*(3), 725-729.

Rogers, M. A., Winter, H., Wolf, C., Heck, M., and Schweizer, J. (1998). Characterization of a 190-kilobase pair domain of human type I hair keratin genes. *Journal of Biological Chemistry, 273*(41), 26683-26691.

Rogers, M. A., Winter, H., Langbein, L., Wolf, C., and Schweizer, J. (2000). Characterization of a 300 kbp region of human DNA containing the type II hair keratin gene domain. *Journal of Investigative Dermatology, 114*(3), 464-472.

Rogers, M. A., Winter, H., Langbein, L., Bleiler, R., and Schweizer, J. (2004a). The human type I keratin gene family: Characterization of new hair follicle specific members and evaluation of the chromosome 17q21.2 gene domain. *Differentiation, 72*(9-10), 527-540.

Rogers, M. A., Langbein, L., Praetzel-Wunder, S., Winter, H., and Schweizer, J. (2006). Human hair keratin-associated proteins (KAPs). *International Review of Cytology, 251*, 209-263.

Rogers, M. A., Langbein, L., Winter, H., Ehmann, C., Praetzel, S., and Schweizer, J. (2002). Characterization of a first domain of human high glycine-tyrosine and high sulfur keratin-associated protein (KAP) genes on chromosome 21q22.1. *Journal of Biological Chemistry, 277*(50), 48993-49002.

Rogers, M. A., Langbein, L., Winter, H., Beckmann, I., Praetzel, S., and Schweizer, J. (2004b). Hair keratin associated proteins: Characterization of a second high sulfur KAP gene domain on human chromosome 21. *Journal of Investigative Dermatology, 122*(1), 147-158.

Rogers, M. A., Edler, L., Winter, H., Langbein, L., Beckmann, I., and Schweizer, J. (2005). Characterization of new members of the human type II keratin gene family and a general evaluation of the keratin gene domain on chromosome 12q13. 13. *Journal of Investigative Dermatology, 124*(3), 536-544.

Rogers, M. A., Langbein, L., Winter, H., Ehmann, C., Praetzel, S., Korn, B., and Schweizer, J. (2001). Characterization of a cluster of human high/ultrahigh sulfur keratin-associated protein genes

embedded in the type I keratin gene domain on chromosome 17q12-21. *Journal of Biological Chemistry, 276*(22), 19440-19451.

Rogers, M. A., Winter, H., Langbein, L., Wollschläger, A., Praetzel-Wunder, S., Jave-Suarez, L. F., and Schweizer, J. (2007). Characterization of human KAP24. 1, a cuticular hair keratin-associated protein with unusual amino-acid composition and repeat structure. *Journal of Investigative Dermatology, 127*(5), 1197-1204.

Roldan, D., Dodero, A., Bidinost, F., Taddeo, H., Allain, D., Poli, M., and Elsen, J. (2010). Merino sheep: a further look at quantitative trait loci for wool production. *Animal, 4*(8), 1330-1340.

Ryder, M. L., and Stephenson, S. K. (1968). *Wool Growth*. London and New York: Academic Press.

Safari, E., Fogarty, N., Gilmour, A., Atkins, K., Mortimer, S., Swan, A., Brien, F., Greeff, J., and Van Der Werf, J. (2007). Genetic correlations among and between wool, growth and reproduction traits in Merino sheep. *Journal of Animal Breeding and Genetics, 124*(2), 65-72.

Schweizer, J., Bowden, P. E., Coulombe, P. A., Langbein, L., Lane, E. B., Magin, T. M., Maltais, L., Omary, M. B., Parry, D. A., and Rogers, M. A. (2006). New consensus nomenclature for mammalian keratins. *The Journal of Cell Biology, 174*(2), 169-174.

Shibuya, K., Obayashi, I., Asakawa, S., Minoshima, S., Kudoh, J., and Shimizu, N. (2004). A cluster of 21 keratin-associated protein genes within introns of another gene on human chromosome 21q22.3. *Genomics, 83*(4), 679-693.

Shimomura, Y., Aoki, N., Schweizer, J., Langbein, L., Rogers, M. A., Winter, H., and Ito, M. (2002). Polymorphisms in the human high sulfur hair keratin-associated protein 1, KAP1, gene family. *Journal of Biological Chemistry, 277*(47), 45493-45501.

Simpson, W. S., and Crawshaw, G. (2002). *Wool: science and technology*. Cambridge: England Woodhead Publishing.

Sparrow, L. G., Dowling, L. M., Loke, V., and Strike, P. M. (1989). Amino acid sequences of wool keratin IF proteins. In G. E. Rogers, P. J. Reis, K. A. Ward & R. C. Marshall (Eds.), *The Biology of Wool and Hair* (pp. 145-155). Netherlands: Springer.

Sparrow, L. G., Robinson, C. P., Caine, J., McMahon, D., and Strike, P. (1992). Type II intermediate-filament proteins from wool. The amino acid sequence of component 5 and comparison with component 7c. *Biochemical Journal, 282(Pt1)*, 291-297.

Srikanth, B., Vaidya, M. M., and Kalraiya, R. D. (2010). O-GlcNAcylation Determines the Solubility, Filament Organization, and Stability of Keratins 8 and 18. *Journal of Biological Chemistry, 285*(44), 34062-34071.

Stein, W. H., and Moore, S. (1948). Chromatography of amino acids on starch columns. Separation of phenylalanine, leucine, isoleucine, methionine, tyrosine, and valine. *Journal of Biological Chemistry, 176*(1), 337-365.

Steinert, P., and Roop, D. (1988). Molecular and cellular biology of intermediate filaments. *Annual Review of Biochemistry, 57*(1), 593-625.

Sumner, R. M. W., and Hawker, H. (1986). Implications of objective measurement for sheep management and wool production. *Proceedings of the New Zealand Society of Animal Production*, *46*, 198-200.

Swart, L., and Haylett, T. (1971). Studies on the high-sulphur proteins of reduced Merino wool. Amino acid sequence of protein SCMKB-IIIB4. *Biochemical Journal, 12(2)*, 201-210.

Swart, L., Haylett, T., and Joubert, F. (1969). Studies on the High-Sulfur Proteins of Reduced Merino Wool Part II: The Isolation of a Homogeneous Protein. *Textile Research Journal, 39*(10), 912-917.

Swart, L. S., and Haylett, T. (1973). Studies on the high-sulphur proteins of reduced Merino wool. Amino acid sequence of protein SCMKB-IIIA3. *Biochemical Journal, 133(4)*, 641-654.

Tatari, Z., Fortier, C., Bobrynina, V., Loiseau, P., Charron, D., and Raffoux, C. (1995). HLA-Cw allele analysis by PCR-restriction fragment length polymorphism: study of known and additional alleles. *Proceedings of the National Academy of Sciences, 92*(19), 8803-8807.

Taylor, P., Atkins, K., and Gilmour, A. (1999). Genetic association between fibre curvature, staple crimp and wool production and quality of Merino sheep. *Proceedings of the Association for the Advancement of Animal Breeding and Genetics, 13*, 456-459.

Turner, H. N. (1964). Relationships between some important characteristics in the Australian Merino. *Wool Technology and Sheep Breeding, 11*(2), 95-101.

Wang, X., Zhao, Z. D., Xu, H. R., Qu, L., Zhao, H. B., Li, T., and Zhang, Z. Y. (2012). Variation and expression of KAP9. 2 gene affecting cashmere trait in goats. *Molecular Biology Reports, 39*(12), 10525-10529.

White, M. B., Carvalho, M., Derse, D., O'Brien, S. J., and Dean, M. (1992). Detecting single base substitutions as heteroduplex polymorphisms. *Genomics, 12*(2), 301-306.

Wilkinson, P. R. (1970). Static electricity: domestic and commercial implications in textile materials. Modern Textiles, *51*(12), 35-40.

Williams, J. (2005). The use of marker-assisted selection in animal breeding and biotechnology. *Revue Scientifique et Technique-Office International des Epizooties, 24*(1), 379.

Wilson, B. W., Edwards, K. J., Sleigh, M. J., Byrne, C. R., and Ward, K. A. (1988). Complete sequence of a type-I microfibrillar wool keratin gene. *Gene, 73*(1), 21-31.

Winter, H., Langbein, L., Krawczak, M., Cooper, D. N., Jave-Suarez, L. F., Rogers, M. A., Praetzel, S., Heidt, P. J., and Schweizer, J. (2001). Human type I hair keratin pseudogene φ hHaA has functional orthologs in the chimpanzee and gorilla: evidence for recent inactivation of the human gene after the Pan-Homo divergence. *Human Genetics, 108*(1), 37-42.

Wood, N. J., Phua, S. H., and Crawford, A. M. (1992). A dinucleotide repeat polymorphism at the glycine-and tyrosine-rich keratin locus in sheep. *Animal Genetics, 23*(4), 391-391.

Wools of New Zealand. (1996). The New Zealand Wool Industry Manual.

Wu, D. D., Irwin, D., and Zhang, Y. P. (2008). Molecular evolution of the keratin associated protein gene family in mammals, role in the evolution of mammalian hair. *BMC Evolutionary Biology, 8*(1), 241.

Wuliji, T., Dodds, K., Andrews, R., and Turner, P. (2011). Selection response to fleece weight, wool characteristics, and heritability estimates in yearling Romney sheep. *Livestock Science, 135*(1), 26-31.

Wuliji, T., Dodds, K., Land, J., Andrews, R., and Turner, P. (2001). Selection for ultrafine Merino sheep in New Zealand: heritability, phenotypic and genetic correlations of live weight, fleece weight and wool characteristics in yearlings. *Animal Science - Glasgow, 72*(2), 241-250.

Yahagi, S., Shibuya, K., Obayashi, I., Masaki, H., Kurata, Y., Kudoh, J., and Shimizu, N. (2004). Identification of two novel clusters of ultrahigh-sulfur keratin-associated protein genes on human chromosome 11. *Biochemical and Biophysical Research Communications, 318*(3), 655-664.

Yu, Z., Gordon, S. W., Nixon, A. J., Bawden, C. S., Rogers, M. A., Wildermoth, J. E., Maqbool, N. J., and Pearson, A. J. (2009). Expression patterns of keratin intermediate filament and keratin associated protein genes in wool follicles. *Differentiation, 77*(3), 307-316.

Yu, Z., Wildermoth, J. E., Wallace, O. A., Gordon, S. W., Maqbool, N. J., Maclean, P. H., Nixon, A. J., and Pearson, A. J. (2011). Annotation of sheep keratin intermediate filament genes and their patterns of expression. *Experimental Dermatology, 20*(7), 582-588.

Zahn, H., Schäfer, K., and Popescu, C. (2005). Wool from animal sources. *Biopolymers Online, 8.*

Zhou, H., Hickford, J. G. H., and Fang, Q. (2006). A two-step procedure for extracting genomic DNA from dried blood spots on filter paper for polymerase chain reaction amplification. *Analytical Biochemistry, 354*(1), 159-161.

Zhou, H., Hickford, J. G. H., Fang, Q., and Lin, Y. S. (2007). Allelic variation of the ovine Toll-like receptor 4 gene. *Developmental & Comparative Immunology, 31*(2), 105-108.

Zhou, H., Gong, H., Yan, W., Luo, Y., and Hickford, J. G. H. (2012). Identification and sequence analysis of the keratin-associated protein 24–1 (KAP24-1) gene homologue in sheep. *Gene, 511*, 62-65.

Zhou, H., Kunhareang, S., Gong, H., Fang, Q., Hu, J., Luo, Y., and Hickford, J. G. H. (2011). Detection of sequence variation and genotyping of polymorphic genes using polymerase chain reaction stem–loop conformational polymorphism analysis. *Analytical Biochemistry, 408*(2), 340-341.

Zhu, M., and Zhao, S. (2007). Candidate gene identification approach: progress and challenges. *International Journal of Biological Sciences, 3*(7), 420-427.

Appendix A

Major Suppliers of Molecular Biology Reagents, Enzymes and Equipment

Bioline, London, UK

Bio-Rad Laboratories, Inc., Hercules, CA, USA

Gilson S.A.S., Villiers Le Bel, French

Integrated DNA Technologies, Coralville, IA, USA

Julabo GmbH, Seelbach, Germany

Lynnon Biosoft, Vaudreuil, Canada

Qiagen GmbH, Hilden, Germany

Quality Scientific Plastics, Inc. San Diego, CA, USA

Quantum Scientific, Queensland, Austrailia

Sigma Chemical Company, St Louis, MO, USA

UVItec Limited, Cambridge, UK

Whatman Bioscience, Middlesex, UK

All other chemicals unless otherwise stated were of AR grade and obtained from **Merck** KGaA, Dermstadt, Germany or **British Drug Houses (BDH)** Chemicals, Poole, England.

Appendix B

Sequence Alignments of *KRTAP* Variants Identified in this Study

(a)*KRTAP1-2*

```
KRTAP1-2*A TAACAACCCTCCTCTCAATCTAACCCCTGACACCATGGCCTGCTGTTCCACCAGTTTCTG     60
KRTAP1-2*B ------------------------------------------------------------     60
KRTAP1-2*C ------------------------------------------------------------     60
KRTAP1-2*D ------------------------------------------------------------     60
KRTAP1-2*E ------------------------------------------------------------     60
KRTAP1-2*F ---------------------------------t--------------------------     60
KRTAP1-2*G ------------------------------------------------------------     60
KRTAP1-2*H ------------------------------------------------------------     60
KRTAP1-2*I ------------------------------------------------------------     60

KRTAP1-2*A TGGATTTCCCATCTGTTCCTCTGTTGGAACCTGTGGCTCCAGCTGCGGCCAGCCAACCTC    120
KRTAP1-2*B ------------------------------------------------------------    120
KRTAP1-2*C ------------------------------------------------------------    120
KRTAP1-2*D ------------------------------------------------------------    120
KRTAP1-2*E ------------------------------------------------------------    120
KRTAP1-2*F ------------------------------------------------------------    120
KRTAP1-2*G ------------------------------------------------------------    120
KRTAP1-2*H ------------------------------------------------------------    120
KRTAP1-2*I ------------------------------------------------------------    120

KRTAP1-2*A CTGCCAGACCAGTTGCTGCCAGCCAACTTCCATCCAAACCAGCTGCTGCCAACCGATCTC    180
KRTAP1-2*B ------------------------------------------------------------    180
KRTAP1-2*C ------------------------------------------------------------    180
KRTAP1-2*D ------------------------------------------------------------    180
KRTAP1-2*E ------------------------------------------------------------    180
KRTAP1-2*F ------------------------------------------------------------    180
KRTAP1-2*G ------------c-----------------------------------------------    180
KRTAP1-2*H ------------------------------------------------------------    180
KRTAP1-2*I ------------------------------------------------------------    180
KRTAP1-2*J ..................------------------------------------------     42
KRTAP1-2*K ..................------------------------------------------     42

KRTAP1-2*A CATCCAGACCAGCTGCTGCCAGCCAACCTGCCTCCAGACCAGTGGCTGTGAGACCGGCTG    240
KRTAP1-2*B ------------------------------------------------------g-----    240
KRTAP1-2*C ------------------------------------------------------a-----    240
KRTAP1-2*D ------------------------------------------------------------    240
KRTAP1-2*E ------------------------------------------------------g-----    240
KRTAP1-2*F ------------------------------------------------------------    240
KRTAP1-2*G ------------------------------------------------------------    240
KRTAP1-2*H ------------------------------------------------------a-----    240
KRTAP1-2*I ------------------------------------------------------a-----    240
KRTAP1-2*J ------------------------------------------------------g-----    102
KRTAP1-2*K ------------------------------------------------------a-----    102

KRTAP1-2*A TGGCATTGGTGGCAGCATTGGCTATGGCCAGGTGGGTAGCAGCGGAGCTGTGAGCAGCCG    300
KRTAP1-2*B ------------------------------------------------------------    300
KRTAP1-2*C ------a-----------------------------------------------------    300
KRTAP1-2*D ------a-----------------------------------------------------    300
KRTAP1-2*E ----------------------------------------t-------------------    300
KRTAP1-2*F ------------------------a-------------------------------------    300
KRTAP1-2*G ------------------------a---------------t-------------------    300
KRTAP1-2*H ------------------------a---------------t-------------------    300
```

```
KRTAP1-2*I ------------------------a---------------------------------    300
KRTAP1-2*J ------------------------------------------------------------    162
KRTAP1-2*K ------------------------a---------------t-----------------    162

KRTAP1-2*A CACCAGGTGGTGCCGCCCTGACTGCCGCGTGGAGGGCACCAGCCTGCCTCCCTGCTGTGT    360
KRTAP1-2*B ------------------------------------------------------------    360
KRTAP1-2*C ------------------------------------------------------------    360
KRTAP1-2*D ------------------------------------------------------------    360
KRTAP1-2*E ------------------------------------------------------------    360
KRTAP1-2*F ------------------------------------------------------------    360
KRTAP1-2*G ------------------------------------------------------------    360
KRTAP1-2*H ------------------------------------------------------------    360
KRTAP1-2*I ------------------------------------------------------------    360
KRTAP1-2*J ------------------------------------------------------------    222
KRTAP1-2*K ------------------------------------------------------------    222

KRTAP1-2*A GGTGAGCTGCACATCCCCGTCCTGCTGCCAGCTGTACTATGCCCAGGCCTCCTGCTGCCG    420
KRTAP1-2*B ------------------------------------------------------------    420
KRTAP1-2*C ------------------------------------------------------------    420
KRTAP1-2*D ------------------------------------------------------------    420
KRTAP1-2*E ------------------------------------------------------------    420
KRTAP1-2*F ------------------------------------------------------------    420
KRTAP1-2*G ------------------------------------------------------------    420
KRTAP1-2*H ------------------------a---------------------------------    420
KRTAP1-2*I ------------------------------------------------------------    420
KRTAP1-2*J ------------------------------------------------------------    282
KRTAP1-2*K ------------------------------------------------------------    282

KRTAP1-2*A CCCATCCTACTGTGGACAGTCCTGCTGCCGCCCAGCCTGCTGCTGCCAGCCCACCTGCAT    480
KRTAP1-2*B ------------------------------------------------------------    480
KRTAP1-2*C ------------------------------------------------------------    480
KRTAP1-2*D ------------------------------------------------------------    480
KRTAP1-2*E ------------------------------------------------------------    480
KRTAP1-2*F ------------------------------------------------------------    480
KRTAP1-2*G ------------------------------------------------------------    480
KRTAP1-2*H ------------------------------------------------------------    480
KRTAP1-2*I ------------------------------------------------------------    480
KRTAP1-2*J ------------------------------------------------------------    342
KRTAP1-2*K ------------------------------------------------------------    342

KRTAP1-2*A TGAGCCCGTCTGTGAGCCCACCTGCTGAAAGCAAGGCTGCTGATTTTCAACTTGAAAGTT    540
KRTAP1-2*B ------------------------------------------------------------    540
KRTAP1-2*C ------------------------------g-----------------------------    540
KRTAP1-2*D ------------------------------------------------------------    540
KRTAP1-2*E --------------------g---------------------------------------    540
KRTAP1-2*F ------------------------------------------------------------    540
KRTAP1-2*G ------------------------------------------------------------    540
KRTAP1-2*H ------------------------------------------------------------    540
KRTAP1-2*I ------------------------------------------------------------    540
KRTAP1-2*J ----------------------------------------------..............    388
KRTAP1-2*K ----------------------------------------------..............    388

KRTAP1-2*A CAACTTCAGTCCATGAA                                             557
KRTAP1-2*B ----------------                                             557
KRTAP1-2*C ----------------                                             557
KRTAP1-2*D ----------------                                             557
KRTAP1-2*E ----------------                                             557
KRTAP1-2*F ----------------                                             557
KRTAP1-2*G ----------------                                             557
KRTAP1-2*H ----------------                                             557
KRTAP1-2*I ----------------                                             557
```

(b) *KRTAP1-4*

```
KRTAP1-4*A ATCCTCCAAGCATTACAATTCTCAGCCCAACTCCTGACACCATGGCCTGTTGCTCCACCA     60
KRTAP1-4*B ------------------------------------------------------------     60
KRTAP1-4*C ------------------------------------------------------------     60
KRTAP1-4*D ------------------------------------------------------------     60
KRTAP1-4*E ------------------------------------------------------------     60
KRTAP1-4*F ------------------------------------------------------------     60
KRTAP1-4*G ------------------------------------------------------------     60
KRTAP1-4*H ------------------------------------------------------------     60
KRTAP1-4*I ------------------------------------------------------------     60

KRTAP1-4*A GCTTCTGTGGATTTCCCACTTGCTCCACTGGTGGGACCTGTGGTTCCAACTTTTGCCAGC    120
KRTAP1-4*B ------------------------------------------------------------    120
KRTAP1-4*C ------------------------------------------------------------    120
KRTAP1-4*D ------------------------------------------------------------    120
KRTAP1-4*E ------------------------------------------------------------    120
KRTAP1-4*F ------------------------------------------------------------    120
KRTAP1-4*G ------------------------------------------------------------    120
KRTAP1-4*H ------------------------------------------------------------    120
KRTAP1-4*I ------------------------------------------------------------    120

KRTAP1-4*A CAACCTGCTGCCAGACCAGCTGCTGCCAGCCAACCTCCATTCAGACCAGCTGCTGCCAGC    180
KRTAP1-4*B ------------------------------------------------------------    180
KRTAP1-4*C ------------------------------------------------------------    180
KRTAP1-4*D ------------------------------------------------------------    180
KRTAP1-4*E ------------------------------------------------------------    180
KRTAP1-4*F ------------------------------------------------------------    180
KRTAP1-4*G ------------------------------------------------------------    180
KRTAP1-4*H ------------------------------------------------------------    180
KRTAP1-4*I ------------------------------------------------------------    180

KRTAP1-4*A CGACCTCTATCCAGACCAGCTGCTGCCAGCCAACTTCCATCCAAACCAGCTGCTGCCAAC    240
KRTAP1-4*B ------------------------------------------------------------    240
KRTAP1-4*C --------c---------------------------------------------------    240
KRTAP1-4*D ------------------------------------------------------------    240
KRTAP1-4*E ------------------------------------------------------------    240
KRTAP1-4*F ----------------------------------------------g-------------    240
KRTAP1-4*G ------------------------------------------------------------    240
KRTAP1-4*H --------c---------------------------------------------------    240
KRTAP1-4*I ------------------------------------------------------------    240

KRTAP1-4*A CGATCTCCATCCAGACCAGCTGCTGCCAGCCAACCTGCCTCCAGACCAGTGGCTGTGAGA    300
KRTAP1-4*B ------------------------------------------------------------    300
KRTAP1-4*C ------------------------------------------------------------    300
KRTAP1-4*D ------------------------------------------------------------    300
KRTAP1-4*E ------------------------------------------------------------    300
KRTAP1-4*F ------------------------------------------------------------    300
KRTAP1-4*G ------------------------------------------------------------    300
KRTAP1-4*H ------------------------------------c-a---------------------    300
KRTAP1-4*I ------------------------------------c-a---------------------    300

KRTAP1-4*A CGGGCTGTGGCATTGGTGGCAGCATTGGCTATGGCCAGGTGGGTAGCAGCGGAGCTGTGA    360
KRTAP1-4*B ------------------------------------------------------------    360
KRTAP1-4*C -t----------a------------a---------------------------------    360
KRTAP1-4*D ------------------------------------------------------------    360
KRTAP1-4*E ------------------------------------------------------------    360
KRTAP1-4*F -c----------------------------------------------------------    360
KRTAP1-4*G -t----------a-----------------------------------------------    360
KRTAP1-4*H -c---------------------c------------------------------------    360
KRTAP1-4*I ------------------------------------------------------------    360
```

```
KRTAP1-4*A GCAGCCGCACCAGGTGGTGCCGCCCTGACTGCCGCGTGGAGGGCACCAGCCTGCCTCCCT    420
KRTAP1-4*B ------------------------------------------------------------    420
KRTAP1-4*C ------------------------------------------------------------    420
KRTAP1-4*D ------------------------------------------------------------    420
KRTAP1-4*E ------------------------------------------------------------    420
KRTAP1-4*F ------------------------------------------------------------    420
KRTAP1-4*G ------------------------------------------------------------    420
KRTAP1-4*H ------------------------------------------------------------    420
KRTAP1-4*I ------------------------------------------------------------    420

KRTAP1-4*A GCTGCGTGGTGAGCTGCACATCCCCGTCCTGCTGCCAGCTGTACTATGCCCAGGCCTCCT    480
KRTAP1-4*B ----t-------------------------------------------------------    480
KRTAP1-4*C ----t-------------------------------------------------------    480
KRTAP1-4*D ------------------------------------------------------------    480
KRTAP1-4*E ----t-------------------------------------------------------    480
KRTAP1-4*F ------------------------------------------------------------    480
KRTAP1-4*G ------------------------------------------------------------    480
KRTAP1-4*H ------------------------------------------------------------    480
KRTAP1-4*I ----t-------------------------------------------------------    480

KRTAP1-4*A GCTGCCGCCCATCCTACTGTGGACAGTCCTGCTGCCGCCCAGCCTGCTGCTGCCAGCCCA    540
KRTAP1-4*B ------------------------------------------------------------    540
KRTAP1-4*C --------------------------------------------------t---------    540
KRTAP1-4*D ------------------------------------------------------------    540
KRTAP1-4*E ------------------------------------------------------------    540
KRTAP1-4*F --------------------------------------------------t---------    540
KRTAP1-4*G ------------------------------------------------------------    540
KRTAP1-4*H ------------------------------------------------------------    540
KRTAP1-4*I ------------------------------------------------------------    540

KRTAP1-4*A CCTGCATTGAGCCCGTCTGTGAGCCCAGCTGCTGTGAGCCCACCTGCTGAAAGCAAGGTT    600
KRTAP1-4*B ---------------------------c--------------------------------    600
KRTAP1-4*C ------------------------------------------------------------    600
KRTAP1-4*D -------------a----------------------------------------------    600
KRTAP1-4*E ------------------------------------------------------------c-    600
KRTAP1-4*F ---------------------------c-------------g------------------    600
KRTAP1-4*G ------------------------------------------------------------    600
KRTAP1-4*H ---------------------------c--------------------------------    600
KRTAP1-4*I ------------------------------------------------------------c-    600

KRTAP1-4*A GCTCATTTAAAATTGCCCAAGACACAGTATCTCTGAAT                         638
KRTAP1-4*B -------------------------------------                         638
KRTAP1-4*C -------------------------------------                         638
KRTAP1-4*D -------------------------------------                         638
KRTAP1-4*E -------------------------------------                         638
KRTAP1-4*F -------------------------------------                         638
KRTAP1-4*G -------------------------------------                         638
KRTAP1-4*H -------------------------------------                         638
KRTAP1-4*I -------------------------------------                         638
```

(c) *KRTAP5-4*

```
KRTAP5-4*A CTGCTCCTCTGACCTACTCCACCCTCAACCCACCAGAACCATGGGCTGCTCTGGCTGTTC     60
KRTAP5-4*B ------------------------------------------------------------     60
KRTAP5-4*C ------------------------------------------------------------     60
KRTAP5-4*D ------------------------------------------------------------     60
KRTAP5-4*E ------------------------------------------------------------     60

KRTAP5-4*A CGGAGGCTGCGGCTCCAGCTGTGGGGGCTGCGGCTCGTGTGGGGGCTGCGGCTCCAGCTG    120
KRTAP5-4*B ------------------------------------------------------------    120
KRTAP5-4*C ------------------------------------------------------------    120
KRTAP5-4*D ----------------------------------------------a-----------    120
KRTAP5-4*E ------------------------------------------------------------    120
```

```
KRTAP5-4*A CTGTGTGCCTGTCTGCTGCTGCAAGCCCGTGTGCTGCTGTGTGCCAGCCTGCTCCTGCTC    180
KRTAP5-4*B ------------------------------------------------------------    180
KRTAP5-4*C ------------------------------------------------------------    180
KRTAP5-4*D ------------------------------------------------------------    180
KRTAP5-4*E ------------------------------------------------------------    180

KRTAP5-4*A CAGCTGTGGCAAAGGGGGCTGCGGCTCCTGTGGGGGCTCCAAGGGGGGCTGCAGCTCCTG    240
KRTAP5-4*B ------------------------------------------------------------    240
KRTAP5-4*C ------------------------------------------------------------    240
KRTAP5-4*D ------------------------------------------------------------    240
KRTAP5-4*E ------------------------------------------------------------    240

KRTAP5-4*A TGGGGGGTCTAAGGGGAGCTGTGGCTCATGTGGAGGCTGTGGCTCCAGCTGCTGCAAGCC    300
KRTAP5-4*B ------------------------------------------------------------    300
KRTAP5-4*C ------------------------------------------------------------    300
KRTAP5-4*D ------------------------------------------------------------    300
KRTAP5-4*E ------------------------------------------------------------    300

KRTAP5-4*A CGTGTGCTGCTGTGTGCCCGTCTGCTCCTGCTCCAGCTGTGGCAAAGGGGGCTGTGGCTC    360
KRTAP5-4*B ------------------------------------------------------------    360
KRTAP5-4*C ------------------------------------------------------------    360
KRTAP5-4*D ------------------------------------------------------------    360
KRTAP5-4*E ------------------------------------------------------------    360

KRTAP5-4*A CAGCTGTGGGGGCTCCAAGGGGGGCTGCGGCTCCTGTGGGGGGTCTAAGGGGGGCTGTGG    420
KRTAP5-4*B ------------------------------------------------------------    420
KRTAP5-4*C ------------------------------------------------------------    420
KRTAP5-4*D ------------------------------------------------------------    420
KRTAP5-4*E ------------------------------------------------------------    420

KRTAP5-4*A CTCGTGTGGGGGCTGTGGCTCTGGCTGCGGCCCCAGCTGCTGTGTGCCCGTGTGCTGCTG    480
KRTAP5-4*B ------------------------------------------------------------    480
KRTAP5-4*C ------------------------------------------------------------    480
KRTAP5-4*D --------------------------------t---------------------------    480
KRTAP5-4*E ------------------------------------------------------------    480

KRTAP5-4*A TGTGCCAGCCTGCTCCTGCTCCAGCTGTGGCAAAGGGGGCTGCGGCTCCTGTG.......    533
KRTAP5-4*B ------------------------------------------------------gctgctc    540
KRTAP5-4*C ---------------------------------------------------.......    533
KRTAP5-4*D ---------------------------------------------------.......    533
KRTAP5-4*E ---------------------------------------------------.......    533

KRTAP5-4*A ......................GCTGCTCCCAGTCCAGCTGCTGCAGACCCTGCTGCTC    570
KRTAP5-4*B ccagtccagctgctgcagaccct-------------------------------------    600
KRTAP5-4*C ......................c-------------------------------------    570
KRTAP5-4*D ......................-------------------------------------    570
KRTAP5-4*E ......................-------------------------------------    570

KRTAP5-4*A CCAGTCCAGCTGCTGCGTCCCCGTGTGCTGCCAGCGCAAGATCTGATGTAACGGTCAGTC    630
KRTAP5-4*B ------------------------------------------------------------    660
KRTAP5-4*C ------------------------------------------------------------    630
KRTAP5-4*D ---------------t--------------------------------------------    630
KRTAP5-4*E ---------------t-------------------a------------------------    630

KRTAP5-4*A CGGCGTTCAGCTAACCTCTTCCAGCCGCGGCATCTGTGAGGCTGTCTG               678
KRTAP5-4*B ------------------------------------------------           708
KRTAP5-4*C ------------------------------------------------           678
KRTAP5-4*D -a----------------------------------------------           678
KRTAP5-4*E ------------------------------------------------           678
```

(d) *KRTAP6*

```
KRTAP6*A   TCTACCCGAGAACAACCTCAACAAGCAACACCATGTGTGGCTACTACGGAAACTACTATG     60
KRTAP6*C   ------------------------------------------------------------     60
KRTAP6*B   ------------------------------------------------------------     60
KRTAP6*D   ------------------------------------------------------------     60
KRTAP6*E   ------------------------c-----------------------------------     60

KRTAP6*A   GCGGCCTCGGCTGTGGAAGCTATGGCTATGGAGGCCTGGGCTGTGGCTATGGCTCCTGCT    120
KRTAP6*C   ----------------------c-------------------------------------    120
KRTAP6*B   ----------------------ca-------------.....................      98
KRTAP6*D   ----------------------ca-------------------------------------    120
KRTAP6*E   ------------------------------------------a-----------gtg---    120

KRTAP6*A   ACGGCTCTGGCTTCCGCAGGCTGGGCTGTGGCTATGGCTGTGGCTATGGCTATGGCTCCC    180
KRTAP6*C   ------------------------------------------------------------    180
KRTAP6*B   .................................-----------------------      123
KRTAP6*D   ------------------------------------------------------------    180
KRTAP6*E   c---t-----------------------------------cct----c----c-----t--    180

KRTAP6*A   GCTCTCTCTGTGGAAGTGGCTATGGCTGTGGCTCCCGCCCTCTCTATGGCTGTGGCTATG    240
KRTAP6*C   ------------------------------------------------------------    240
KRTAP6*B   ---------------------------a----------t------g---aa---------    183
KRTAP6*D   ---------------------------a----------t------g---aa---------    240
KRTAP6*E   -tagg--g.....................----gtg--ta-gg--g-----a--------    219

KRTAP6*A   GATGTGGCTCTGGCTATGGCTCTGGCTTTGGCTACTACTATTGAGGACGCCACAGAAGAC    300
KRTAP6*C   ------------------------------------------------------------    300
KRTAP6*B   ----c-----------------------------------------t-----g------    243
KRTAP6*D   ----c-----------------------------------------t-----ga-----    300
KRTAP6*E   ----c......-----------------------------------a--t----aga-----    273

KRTAP6*A   TCTCATCCTCTATACCT                                               317
KRTAP6*C   -----------------                                               317
KRTAP6*B   -----------------                                               260
KRTAP6*D   -----------------                                               317
KRTAP6*E   -----------------                                               290
```

(e) *KRTAP7-1*

```
KRTAP7-1*A ACTTGCTCTTCACATTCTATCCAAATCCTTCCCACTCCTGCCACAATGACTCGTTTCTTT     60
KRTAP7-1*B ------------------------------------------------------------     60

KRTAP7-1*A TGCTGCGGAAGCTACTTCCCAGGCTATCCTTCCTATGGAACCAATTTCCACAGGACCTTC    120
KRTAP7-1*B ------------------------------------------------------------    120

KRTAP7-1*A AGAGCCACCCCCCTGAACTGCGTTGTGCCCCTTGGCTCTCCCCTTGGTTATGGATGCAAT    180
KRTAP7-1*B ------------------------------------------------------------    180

KRTAP7-1*A GGCTACAGCTCCCTGGGCTACGGTTTCGGTGGAAGCAGCTTTAGCAACCTGGGCTGTGGC    240
KRTAP7-1*B ------------------------------------a-----------------------    240

KRTAP7-1*A TATGGGGGCAGCTTTTATAGGCCATGGGGCTCTGGCTCTGGCTTTGGCTACAGCACCTAC    300
KRTAP7-1*B ------------------------------------------------------------    300

KRTAP7-1*A TGATGGACCATGGCTCCAGATGACTAC                                     327
KRTAP7-1*B ---------------------------                                     327
```

(f) *KRTAP8-1*

```
KRTAP8-1*A CATTCCCTGCTCTCCAAGCCGCCCAACCCAGACACCATGAGCTACTGCTTCTCCAGCACC     60
KRTAP8-1*B -------------------t----------------------------------------     60
KRTAP8-1*C ------------------------------------------------------------     60
```

```
KRTAP8-1*D ------------------------------------------------------------     60
KRTAP8-1*E ------------------------------------------------------------     60

KRTAP8-1*A GTCTTCCCAGGTTGCTACTGGGGCAGCTATGGCTACCCGCTGGGCTACAGTGTGGGCTGT    120
KRTAP8-1*B ------------------------------------------------------------    120
KRTAP8-1*C -------------t----------------------------------------------    120
KRTAP8-1*D ------------------------------------------------------------    120
KRTAP8-1*E ------------------------------------------------------------    120

KRTAP8-1*A GGCTACGGTAGTACCTACTCCCCAGTGGGCTATGGCTTCGGCTATGGCTACAACGGCTCT    180
KRTAP8-1*B ------------------------------------------------------------    180
KRTAP8-1*C ------------------------------------------------------------    180
KRTAP8-1*D ---------------a--------------------------------------------    180
KRTAP8-1*E ---------------------------------c--------------------------    180

KRTAP8-1*A GGGGCCTTCGGTTGCCGAAGATTCTGGCCATTTGCTCTCTACTGATTTGCTGAAATACCA    240
KRTAP8-1*B ------------------------------------------------------------    240
KRTAP8-1*C ------------------------------------------------------------    240
KRTAP8-1*D ------------------------------------------------------------    240
KRTAP8-1*E ------------------------------------------------------------    240

KRTAP8-1*A GAGGCATGGAATCTTCTC                                             258
KRTAP8-1*B ------------------                                             258
KRTAP8-1*C ------------------                                             258
KRTAP8-1*D ------------------                                             258
KRTAP8-1*E ------------------                                             258
```

(g) *KRTAP8-2*

```
KRTAP8-2*A TAGGCAGTCAGTCATCCTGAAACAATTTTAAAGAGTAATGAAGAAGCCTAGGGTGGGTAT     60
KRTAP8-2*B ------------------------------------------------------------     60

KRTAP8-2*A TTGTTGCCACACCCCTTAGTGGCAGTGTATAAAAGGCTTGGACAGATGGAGCAAGTTATT    120
KRTAP8-2*B ------t-----------------------------------------------------    120

KRTAP8-2*A CCCAGGGAATGGGGTCTTTGCCGCTGAACACTCATCTCTTATCAAAAATGTCCTGTGGCT    180
KRTAP8-2*B ------------------------------------------------------------    180

KRTAP8-2*A TTTTCAATGAAGGCATCTACCCAGGTTACTACTGGGGCAGTTGGGGATACCCCCTGGGCT    240
KRTAP8-2*B ------------------------------------------------------------    240

KRTAP8-2*A ACAGTGTTGGCTGTGGATATGGTAGCACCTATTCCCCAGTGGGCTATGGGTTCGGATACG    300
KRTAP8-2*B ------------------------------------------------------------    300

KRTAP8-2*A GATATGGTGGCTTCCGCCCATTTGACTACAGAAGATACTGGACATTTGACCTCTATTAAT    360
KRTAP8-2*B ------------------------------------------------------------    360

KRTAP8-2*A CAACTTCTAACCTCATGAACTGCACAATAATCTGCATCCCTCAGAACCAAACTTGAAGCA    420
KRTAP8-2*B ------------------------------------------------------------    420

KRTAP8-2*A ATATTGTACTGACAGCTTTCCAGAGTACAGAACGTGGACTTCATATTCTCTAT           473
KRTAP8-2*B -----------------------------------------------------           473
```

(h) *KRTAP11-1*

```
KRTAP11-1*A TGCATCTCTCAACCAGCACCATGTCCTACAGCTGCTCCACAAGGAACTGCTCTTCCAGGC     60
KRTAP11-1*B ------------------------------------------------------------     60
KRTAP11-1*C ------------------------------------------------------------     60
RTKAP11-1*D ------------------------------------------------------------     60
KRTAP11-1*E ------------------------------------------------------------     60
KRTAP11-1*F ------------------------------------------------------------     60
```

```
KRTAP11-1*A GGATTGGAGGAGAATACACTGTTCCAGTGGTCACAGTTTCTTCCCCGGATGCCGATTGCC    120
KRTAP11-1*B ------------------------------------------------------t-------    120
KRTAP11-1*C ----------------------------t----------------------t-------    120
RTKAP11-1*D ------------------------------------------------------------    120
KRTAP11-1*E ------------------------------------------------------t-------    120
KRTAP11-1*F ------------------------------------------------------------    120

KRTAP11-1*A TGAGTGGCATCTATTTGCCCAGCTCCTTCCAAACGGGCTCCTGGCTCCTGGACCACTGTC    180
KRTAP11-1*B ------------------------------------------------------------    180
KRTAP11-1*C ------------------------------------------------------------    180
RTKAP11-1*D ------------------------------------------------------------    180
KRTAP11-1*E ------------------------------------------------------------    180
KRTAP11-1*F ------------------------------------------------------------    180

KRTAP11-1*A AGGAGACCTGCTGTGAGCCCACTGTTTGCCAGTCAACTTGCTACCAGCCAACTCCTTGTG    240
KRTAP11-1*B ------------------------------------------------------------    240
KRTAP11-1*C ------------------------------------------------------------    240
RTKAP11-1*D ------------------------------------------------------------    240
KRTAP11-1*E ------------------------------------------------------------    240
KRTAP11-1*F ------------------------------------------------------------    240

KRTAP11-1*A TCTCCAGCCCTGTGCGGGTGACCTCTCGGCAAACCACCTGTGTCTCCAGTCCCTGTTCGA    300
KRTAP11-1*B ------------------------------------------------------------    300
KRTAP11-1*C ------------------------------------------------------------    300
RTKAP11-1*D ------------------------------------------------------------    300
KRTAP11-1*E ------------------------------------------------------------    300
KRTAP11-1*F ------------------------------------------------------------    300

KRTAP11-1*A CTACCTGCAGCCGGCCACTCACCTTTATCTCCAGTGGCTGTCAGCCTCTGAGTGGCGTCT    360
KRTAP11-1*B ------------------------------------------------------------    360
KRTAP11-1*C ------------------------------------------------------------    360
RTKAP11-1*D ------------------------------------------------------------    360
KRTAP11-1*E ----g-------------------------------------------------g---------    360
KRTAP11-1*F ----a-------------------------------------------------------    360

KRTAP11-1*A CTACTGTGTGCAAGCCAGTGAGAAGCATCTCCACTGTCTGCCAACCGGTGGGAGGAGTCT    420
KRTAP11-1*B ------------------------------------------------------------    420
KRTAP11-1*C ------------------------------------------------------------    420
RTKAP11-1*D ------------------------------------------------------------    420
KRTAP11-1*E ------------------------------------------------------------    420
KRTAP11-1*F ------------------------------------------------------------    420

KRTAP11-1*A CCACCATCTGCCAACCTACCTGCGGGGTCTCCAGGACGTACCAGCAGTCCTGCGTGTCCA    480
KRTAP11-1*B ------------------------------------------------------------    480
KRTAP11-1*C ------------------------------------------------------------    480
RTKAP11-1*D ----------------------t-------------------------------------    480
KRTAP11-1*E ------------------------------------------------------------    480
KRTAP11-1*F ------------------------------------------------------------    480

KRTAP11-1*A GCTGCAGAAGAATTTGCTAAGTTCAGAAGCCCATGAGTGAATCAAGATTCCA            532
KRTAP11-1*B ------------------------------------------------            532
KRTAP11-1*C ------------------------------------------------            532
RTKAP11-1*D ------------------------------------------------            532
KRTAP11-1*E ------------------------------------------------            532
KRTAP11-1*F ------------------------------------------------            532
```

(i) *KRTAP13-3*

```
KRTAP13-3*A TACATTCAAACTCAGAATCTTCTCACTGTAGCTCAGCTGAACTCACATCTCCTGTGAACA     60
KRTAP13-3*B ------------------------------------------------------------     60
KRTAP13-3*C ------------------------------------------------------------     60
KRTAP13-3*D ------------------------------------------------------------     60
KRTAP13-3*E ------------------------------------------------------------     60
```

```
KRTAP13-3*A TGTCCTACAACTGTTGTTCCAGAAACTTCTCCTCCTGCTCCCTTGGGGGCCACCTGCGCT    120
KRTAP13-3*B ------------------------------------------------------------    120
KRTAP13-3*C ------------------------------------------------------------    120
KRTAP13-3*D ------------------------------------------------------------    120
KRTAP13-3*E ------------------------------------------------------------    120

KRTAP13-3*A ACTCAGGCTCCTCCTGTGGCTCCTCCTTCCCCAGAAACCTGGTCTACAGCACTGATCTCT    180
KRTAP13-3*B ------------------------------------------------------------    180
KRTAP13-3*C ------------------------------------------------------------    180
KRTAP13-3*D ------------------------------------------------------------    180
KRTAP13-3*E ------------------------------------------------------------    180

KRTAP13-3*A GCCCTCGAAGCTCCTGCCAGCTGGGCTCTTCTCTCTACAGTCGAGAGACCTGCTGTGTGC    240
KRTAP13-3*B ------------------------------------------------------------    240
KRTAP13-3*C ------------------------------------------------------------    240
KRTAP13-3*D ------------------------------------------------------------    240
KRTAP13-3*E ------------------------------------------------------------    240

KRTAP13-3*A CCATCAGGACCCAGACGTTCCGTGTGGTGTCCCGTCCCTGCCAGACGTCCTGCCGCCGTC    300
KRTAP13-3*B ------------------------------------------------------------    300
KRTAP13-3*C ------------------------------------------------------------    300
KRTAP13-3*D ---------------------------------------------------------t--    300
KRTAP13-3*E ------------------------------t-----------------------------    300

KRTAP13-3*A GGAGGACCTCCACCTTCTCCAGTCCCTGCAAGACGACTCACCATGGATCTCTGGGCTGTA    360
KRTAP13-3*B a----------------------------------------------------------    360
KRTAP13-3*C a----------------------------------------------------------    360
KRTAP13-3*D a----------------------------------------------------------    360
KRTAP13-3*E ------------------------------------------------------------    360

KRTAP13-3*A AGTCCAGCAGCTGCAGCTCCCTGAGCTCTGGATCCAGAAGGTGCTACTCAGTGGGCTGTG    420
KRTAP13-3*B ------------------------------------------------------------    420
KRTAP13-3*C ------------------------------------------------------------    420
KRTAP13-3*D ------------------------------------------------------------    420
KRTAP13-3*E ------------------------------------------------------------    420

KRTAP13-3*A GATCCCGTGTCTTCAGACCCCTGGGCTATGGAGTCTGTGGCTTCCCTTCCCTGGGCTGTG    480
KRTAP13-3*B ------------------------------------------------------------    480
KRTAP13-3*C -------------------------c----------------------------------    480
KRTAP13-3*D ------------------------------------------------------------    480
KRTAP13-3*E ------------------------------------------------------------    480

KRTAP13-3*A GATCCCGATTCTGGCACCCAATTAATTTTCCCTGCAGAAGTTTCCATTAATCTTGCTACT    540
KRTAP13-3*B ------------------------------------------------------------    540
KRTAP13-3*C ------------------------------------------------------------    540
KRTAP13-3*D ------------------------------------------------------------    540
KRTAP13-3*E ------------------------------------------------------------    540

KRTAP13-3*A TACCAACCTGTAGATCTGGCTTCTACTGATTAACTTGTAGAAGAGCCAAATTCA          594
KRTAP13-3*B ----------------------------------------------------          594
KRTAP13-3*C ----------------------------------------------------          594
KRTAP13-3*D ----------------------------------------------------          594
KRTAP13-3*E ----------------------------------------------------          594
```

(j) *KRTAP24-1*

```
KRTAP24-1*A AAATGTTGCCATGTTATTGTCCCAGGCAGTGAGAAGGAAGAGAGTCTGGGGAGGATCCAG     60
KRTAP24-1*B ------------------------------------------------------------     60
KRTAP24-1*C ------------------------------------------------------------     60
KRTAP24-1*D ------------------------------------------------------------     60
```

```
KRTAP24-1*A ACAATCCCCAGGGAAGATATAAAAGGGCGTCAAGCCACAGTCTCGCCATACCTGCACGGC    120
KRTAP24-1*B ------------------------------------------------------------    120
KRTAP24-1*C ------------------------------------------------------------    120
KRTAP24-1*D ------------------------------------------------------------    120

KRTAP24-1*A CATCACTGAACATACACCCAGGCTCCATGGCTTTTCTAGGCTATCCTGGAAACTGTAGTG    180
KRTAP24-1*B ------------------------------------------------------------    180
KRTAP24-1*C ------------------------------------------------------------    180
KRTAP24-1*D ------------------------------------------------------------    180

KRTAP24-1*A GCGTATCCTACAGAACTCACTATTATTTCCCAGTGACTGGTTCTGTTGCTCTTTGCTCCA    240
KRTAP24-1*B ------------------------------------------------------------    240
KRTAP24-1*C ------------------------------------------------------------    240
KRTAP24-1*D ------------------------------------------------------------    240

KRTAP24-1*A GACATGTAAGCCCTACGTTTGGGCTCAGCCTACCCAGTAGCTACCATGGGAATCTCTGGC    300
KRTAP24-1*B ------------------------------------------------------------    300
KRTAP24-1*C ------------------------------------------------------------    300
KRTAP24-1*D ---------------a--------------------------------------------    300

KRTAP24-1*A TCCTGGATAACTGCCAAGAAACTTGTGGAGAAGCCCCGACCTGTGAATCTCCCTGTTCTG    360
KRTAP24-1*B ----------------------------------t-----------------------    360
KRTAP24-1*C ----------------------------------t--t--------------------    360
KRTAP24-1*D ------------------------------------------------------------    360

KRTAP24-1*A AGCCCAAGACCTGCACCACAACTTGTGACCAATCAAACTCGTCTGTGCCCTGCAACTCTC    420
KRTAP24-1*B ------------------------------------------------------------    420
KRTAP24-1*C ------------------------------------------------------------    420
KRTAP24-1*D ------------------------------------------------------------    420

KRTAP24-1*A CAACAGGGGGCCAAATCTGCAGTGCCCGTGAAACGACCAACATCGGACCCAGCCTCAGCT    480
KRTAP24-1*B ------------------------------------------------------------    480
KRTAP24-1*C ------------------------------------------------------------    480
KRTAP24-1*D ----------------------------------------------------------t---    480

KRTAP24-1*A GCAATCAGTGCCCTCAGACCAAGGGGTATGTATCTGATGGCTGCACCCCCAGCCGACATA    540
KRTAP24-1*B ------------------------------------------------------------    540
KRTAP24-1*C ------------------------------------------------------------    540
KRTAP24-1*D ------------------------------------------------------------    540

KRTAP24-1*A CATCCAAAGCTTGCCAGACCCTCGGCAATGGCTTTAAATGCTTTGGGCAACTTAACTGCT    600
KRTAP24-1*B ------------------------------------------------------------    600
KRTAP24-1*C ------------------------------------------------------------    600
KRTAP24-1*D ------------------------------------------------------------    600

KRTAP24-1*A TATCCAAGAGTTTCCAGCCCCTAAGCCACTACAGACTGGGCAGTTTTGGATACAGAAGCT    660
KRTAP24-1*B ------------------------------------------------------------    660
KRTAP24-1*C ----------------------------g-------------------------------    660
KRTAP24-1*D ------------------------------------------------------------    660

KRTAP24-1*A ACCAAGATCTTGGCTTCATACCCAGTGGCTTCTCGGCATCACGATATATCACCAACAGCT    720
KRTAP24-1*B ------------------------------------------------------------    720
KRTAP24-1*C ------------a-----------------------------------------------    720
KRTAP24-1*D ------------------------------------------------------------    720

KRTAP24-1*A GCCAACGCCAAAACTATTTAATAAGAAATAGCCAATGTCCATATGATTGGCATAGGAGAT    780
KRTAP24-1*B ------------------------------------------------------------    780
KRTAP24-1*C ------------------------------------------------------------    780
KRTAP24-1*D ------------------------------------------------------------    780

KRTAP24-1*A GCCCTCCACTGAGCTGTTTTGCAAGAAACTTCCGGTCTCTAAGCTCTATACCAAGTTCCT    840
KRTAP24-1*B ------------------------------------------------------------    840
KRTAP24-1*C ------------------------------------------------------------    840
KRTAP24-1*D ------------------------------------------------------------    840
```

```
KRTAP24-1*A TCCCTCCTCTGAGGTATTTGTATGGTGGTTACAGACCTCTGAATTGCTACCGATCAACTT    900
KRTAP24-1*B ------------------a----------------------------------------    900
KRTAP24-1*C ------------------a----------------------------------------    900
KRTAP24-1*D ------------------a----------------------------------------    900

KRTAP24-1*A ATTGAAATTATAGCTGCTAGAAATTGTCCCAAGGTGCAGGTGCCACCTTCTCTAATAATC    960
KRTAP24-1*B ------------------------------------------------------------    960
KRTAP24-1*C ------------------------------------------------------------    960
KRTAP24-1*D ------------------------------------------------------------    960

KRTAP24-1*A TGCTGAATTACTTCATTTGTC                                          981
KRTAP24-1*B ---------------------                                          981
KRTAP24-1*C ---------------------                                          981
KRTAP24-1*D ---------------------                                          981
```

Appendix C

Genotypes of Two *KRTAPs* and Their Wool Trait Measurement Data

1. *KRTAP1-2* genotype and wool trait data

Year	LambID	SireID	EweID	DOB	Birth rank	Sex	*KRTAP1-2*	GFW	CFW	Yield	MFD	MSL	FDSD	MSS	CVFD	PF	MFC
2005	201	Stoneyhurst	384/02	30/09/2005	1	E	*AF*	2.64	2.0	75.2	18.1	86	3.7	34	20.6	1.17	93.1
2005	202	MV144	105	1/10/2005	1	R	*BF*	3.00	2.2	73.4	17.2	86	3.9	24	22.5	1.00	88.2
2005	204	Stoneyhurst	370/02	1/10/2005	2	R	*AG*	2.34	1.6	70.6	19.0	93	3.8	26	20.3	1.17	81.5
2005	205	Stoneyhurst	389/02	2/10/2005	1	E	*AG*	2.82	1.9	66.5	20.9	80	5.0	18	23.7	5.25	82.9
2005	207	Stoneyhurst	9	2/10/2005	1	R	*AH*	3.62	2.5	68.3	19.3	60	3.6	40	18.5	1.67	85.6
2005	210	MV144	133	3/10/2005	2	R	*FG*	2.62	1.9	73.6	16.1	75	3.8	26	23.5	0.25	92.7
2005	216	MV144	85	4/10/2005	2	.	*BF*	2.86	2.3	83.7	17.2	91	5.6	23	32.6	2.58	88.8
2005	217	Stoneyhurst	16	4/10/2005	1	.	*AA*	2.50	1.8	73.4	20.7	76	4.7	45	22.7	4.50	87.8
2005	218	MV144	104	2/10/2005	1	R	*GJ*	2.66	2.1	81.8	17.2	68	3.1	24	18.0	0.50	73.8
2005	219	MV144	97	28/09/2005	1	E	*BB*	2.22	1.6	74.5	18.3	76	4.0	21	21.6	1.33	101.0
2005	221	Stoneyhurst	150	4/10/2005	2	.	*AH*	3.32	2.2	68.1	24.7	94	7.1	14	28.8	22.58	90.3
2005	223	MV144	129	4/10/2005	2	E	*BH*	2.80	2.3	80.5	15.5	94	3.4	38	21.7	0.50	64.3
2005	224	MV144	110	4/10/2005	1	E	*GH*	2.82	2.1	76.3	17.4	75	3.5	54	20.1	0.83	88.3
2005	225	MV144	66	5/10/2005	1	R	*AG*	2.84	2.3	81.6	18.8	84	3.4	26	18.2	0.67	89.2
2005	227	Stoneyhurst	369/02	5/10/2005	2	R	*FH*	2.38	1.9	67.2	20.2	84	4.6	31	23.0	3.33	96.1
2005	229	MV144	64	4/10/2005	2	R	*BF*	2.04	1.5	76.0	16.8	92	3.1	26	18.2	0.50	87.5
2005	230	MV144	64	4/10/2005	2	E	*FG*	1.90	1.5	78.4	18.0	82	4.0	21	22.1	1.08	80.3
2005	231	MV144	121	7/10/2005	1	E	*FG*	2.50	2.0	78.2	16.3	81	2.8	32	17.2	0.00	93.8
2005	232	MV144	53	7/10/2005	1	R	*BG*	2.72	2.0	74.6	14.6	71	3.8	28	25.8	0.17	95.3
2005	234	Stoneyhurst	40	7/10/2005	1	E	*BH*	3.12	2.1	68.9	23.4	91	5.2	28	22.1	9.92	84.3
2005	235	Stoneyhurst	1	7/10/2005	1	E	*AH*	2.38	1.6	70.0	21.4	73	3.6	38	16.8	1.08	119.9
2005	236	MV144	134	7/10/2005	1	E	*BH*	2.04	1.5	77.4	17.0	88	3.2	18	18.8	0.17	86.6
2005	237	MV144	79	8/10/2005	1	R	*BF*	2.94	2.3	77.9	17.7	71	3.1	36	17.8	0.17	87.9
2005	238	MV144	80	8/10/2005	1	R	*BG*	2.88	2.3	81.3	15.7	72	2.9	54	18.5	0.08	92.1
2005	239	Stoneyhurst	145	7/10/2005	1	E	*AF*	2.48	1.5	63.2	19.8	69	3.8	17	19.3	0.92	82.1
2005	241	Stoneyhurst	380/02	7/10/2005	2	E	*AG*	2.10	1.5	70.3	17.1	61	3.9	32	22.7	0.42	105.2
2005	242	Stoneyhurst	392/02	7/10/2005	2	R	*AB*	2.50	1.7	67.7	21.3	80	5.4	20	25.1	5.17	96.8
2005	243	Stoneyhurst	392/02	7/10/2005	2	E	*BH*	2.30	1.6	68.0	21.3	82	4.3	31	20.0	4.00	87.7

2005	244	MV144	62	6/10/2005	1	R	*FG*	2.76	2.2	79.8	14.9	72	2.6	36	17.5	0.50	88.5
2005	245	Stoneyhurst	12	6/10/2005	2	R	*FH*	2.42	1.8	74.7	20.6	81	4.1	45	20.0	1.58	82.6
2005	247	Stoneyhurst	296/02	8/10/2005	2	E	*AG*	2.08	1.3	67.2	16.3	68	3.1	26	19.0	0.08	82.1
2005	248	Stoneyhurst	296/02	8/10/2005	2	E	*AG*	2.52	1.9	75.6	20.2	69	3.7	45	18.5	1.25	82.3
2005	250	MV144	72	18/10/2005	2	E	*AG*	2.38	1.7	74.3	17.6	80	4.0	33	23.0	1.08	88.0
2005	251	Stoneyhurst	297/02	9/10/2005	1	R	*BH*	2.76	2.0	72.4	19.3	90	2.7	39	13.9	0.00	89.5
2005	253	MV144	99	9/10/2005	1	R	*GG*	2.98	2.3	79.0	16.3	83	3.3	33	19.9	0.17	78.9
2005	254	Stoneyhurst	358/02	9/10/2005	1	R	*BH*	2.84	2.1	76.0	18.2	82	3.1	29	17.1	0.17	99.4
2005	255	MV144	138	9/10/2005	1	E	*BG*	2.60	2.0	75.5	16.0	79	2.8	36	17.6	0.42	95.5
2005	256	MV144	141	9/10/2005	1	R	*AB*	2.74	1.9	68.8	16.4	80	3.3	34	20.3	0.50	98.4
2005	257	Stoneyhurst	376/02	9/10/2005	2	R	*FH*	2.26	1.5	67.9	20.0	78	5.7	13	28.3	6.33	101.8
2005	258	Stoneyhurst	376/02	9/10/2005	2	E	*AF*	2.48	1.9	77.4	18.0	76	4.2	16	23.0	1.08	92.6
2005	259	Stoneyhurst	2	9/10/2005	1	E	*AF*	2.86	2.1	74.5	20.4	96	5.5	32	26.7	7.17	79.2
2005	260	Stoneyhurst	373/02	10/10/2005	1	R	*AG*	2.62	1.9	74.3	17.0	69	3.1	31	18.1	0.17	77.6
2005	261	MV144	108	10/10/2005	1	E	*GH*	1.60	1.3	80.0	18.7	66	3.7	31	19.6	1.00	106.8
2005	262	Stoneyhurst	10	10/10/2005	2	R	*AF*	2.08	1.6	78.1	18.1	68	3.0	26	16.3	0.08	103.9
2005	263	Stoneyhurst	10	10/10/2005	2	E	*HH*	1.58	1.1	70.2	21.6	62	5.1	33	23.7	6.75	93.1
2005	264	Stoneyhurst	386/02	10/10/2005	2	.	*AG*	2.32	1.6	69.8	20.7	83	4.1	26	19.7	1.25	88.3
2005	265	Stoneyhurst	386/02	10/10/2005	2	.	*AA*	1.66	1.1	71.4	16.1	67	3.4	11	21.1	0.33	84.5
2005	267	Stoneyhurst	11	10/10/2005	1	E	*BH*	2.26	1.5	67.2	20.8	83	4.3	30	20.7	1.83	96.6
2005	273	Stoneyhurst	14	11/10/2005	2	E	*GH*	2.32	1.7	71.8	21.1	91	4.6	41	21.7	4.00	90.7
2005	274	MV144	117	11/10/2005	2	R	*BH*	2.48	1.8	74.1	17.2	84	4.4	32	25.4	0.83	84.2
2005	279	Stoneyhurst	360/02	12/10/2005	2	E	*AG*	2.44	1.6	68.3	20.9	90	4.9	38	23.6	3.50	101.2
2005	281	Stoneyhurst	388/02	12/10/2005	2	E	*AA*	1.64	1.0	64.7	19.0	68	4.1	29	21.5	0.83	108.5
2005	282	Stoneyhurst	291/02	12/10/2005	2	R	*AH*	1.92	1.3	69.2	16.2	70	3.6	15	22.2	0.67	97.0
2005	283	Stoneyhurst	291/02	12/10/2005	2	R	*AH*	2.28	1.6	70.6	17.4	64	3.3	50	18.8	0.58	96.9
2005	284	Stoneyhurst	356/02	12/10/2005	2	E	*AF*	2.52	1.7	69.3	19.6	81	4.2	30	21.2	1.75	93.7
2005	285	Stoneyhurst	356/02	12/10/2005	2	E	*AG*	2.36	1.6	68.7	20.8	88	4.6	26	21.9	3.75	81.6
2005	286	Stoneyhurst	299/02	12/10/2005	1	R	*AG*	3.14	2.6	83.7	20.1	82	5.3	45	26.2	5.33	87.4
2005	288	MV144	107	12/10/2005	1	E	*GG*	2.68	1.7	66.7	16.0	73	3.1	23	19.5	0.17	92.4
2005	289	MV144	113	12/10/2005	1	E	*AB*	2.40	2.0	83.0	17.0	78	3.3	33	19.5	0.58	78.2
2005	290	MV144	124	12/10/2005	2	E	*GG*	2.14	1.6	76.1	18.6	83	3.7	31	20.1	0.58	98.3
2005	291	MV144	124	12/10/2005	2	R	*AG*	2.22	1.7	76.9	16.2	69	3.5	35	21.7	0.67	84.7
2005	292	MV144	390/02	10/10/2005	1	R	*FG*	2.48	1.8	75.0	15.9	68	3.0	37	19.2	0.42	97.9
2005	294	Stoneyhurst	17	13/10/2005	2	R	*AG*	2.38	1.4	62.1	18.5	72	3.5	23	19.1	1.08	103.9
2005	296	MV144	131	14/10/2005	1	E	*GG*	1.88	1.4	75.4	16.0	67	3.4	26	21.3	0.33	92.1
2005	298	MV144	87	14/10/2005	2	E	*FG*	2.00	1.5	74.1	16.9	74	3.4	14	20.5	0.58	93.8
2005	299	Stoneyhurst	30	14/10/2005	2	.	*AH*	1.98	1.2	63.0	17.9	76	3.8	11	21.4	0.67	100.7
2005	300	Stoneyhurst	30	14/10/2005	2	.	*AH*	1.82	1.2	68.8	19.3	75	3.9	22	20.2	1.33	101.9
2005	301	Stoneyhurst	28	14/10/2005	1	R	*AH*	2.80	2.0	70.0	19.7	80	3.6	24	18.5	0.92	91.8

2005	303	Stoneyhurst	364/02	14/10/2005	1	R	*GH*	2.60	1.7	67.3	17.7	63	4.6	32	25.8	0.83	97.9
2005	304	MV144	65	14/10/2005	1	R	*BH*	2.28	1.8	81.3	16.3	89	3.0	12	18.5	0.17	91.0
2005	307	MV144	52	15/10/2005	1	E	*FG*	2.22	1.7	75.9	16.2	68	3.8	27	23.2	0.67	80.5
2005	309	MV144	11	15/10/2005	1	R	*BF*	2.80	2.1	76.1	18.8	76	3.8	40	20.2	0.92	94.3
2005	310	MV144	132	15/10/2005	1	R	*BG*	2.46	1.9	78.7	18.3	82	3.1	49	17.0	0.58	96.6
2005	311	MV144	135	15/10/2005	1	E	*AG*	2.82	2.2	80.2	16.5	83	5.6	27	33.9	1.33	64.7
2005	313	MV144	136	15/10/2005	1	E	*FG*	2.28	1.7	75.2	15.8	64	3.9	31	24.5	0.42	85.5
2005	314	MV144	136	15/10/2005	2	E	*GH*	1.86	1.4	78.3	19.0	70	3.7	33	19.4	0.42	101.8
2005	316	Stoneyhurst	26	15/10/2005	1	E	*AF*	1.96	1.3	67.3	18.7	79	4.3	21	23.1	1.58	102.1
2005	317	Stoneyhurst	49	14/10/2005	2	R	*AA*	2.54	1.6	65.7	17.5	98	3.1	35	17.9	0.17	83.7
2005	321	MV144	NT	38642	1	.	*BF*	2.42	1.8	75.6	18.6	94	3.6	30	19.5	0.75	97.6
2005	322	Stoneyhurst	148	17/10/2005	1	E	*AF*	2.86	1.9	67.1	18.5	86	4.2	32	22.5	1.58	100.9
2005	323	MV144	391/02	18/10/2005	1	R	*FG*	2.44	1.8	73.4	18.2	69	3.0	42	16.4	0.08	94.4
2005	324	Stoneyhurst	368/02	18/10/2005	1	E	*AA*	2.12	1.6	76.4	17.6	93	3.3	36	18.7	0.25	82.5
2005	325	MV144	126	18/10/2005	1	E	*AB*	2.72	2.2	83.0	16.8	74	3.6	46	21.4	0.50	83.7
2005	327	MV144	70	18/10/2005	1	E	*BH*	2.66	2.0	78.6	17.6	70	3.2	50	18.2	0.42	82.1
2005	329	MV144	75	18/10/2005	2	E	*BF*	2.00	1.5	76.9	18.8	92	3.5	37	18.5	0.58	85.8
2005	331	Stoneyhurst	38	18/10/2005	1	.	*AA*	2.38	1.7	71.9	19.4	88	3.7	16	19.2	0.83	83.5
2005	332	MV144	19	18/10/2005	1	E	*GG*	2.88	2.3	82.4	16.3	82	2.9	33	17.6	0.17	74.8
2005	334	Stoneyhurst	5	19/10/2005	2	E	*AH*	1.88	1.2	66.7	18.0	82	3.7	21	20.3	1.08	95.8
2005	335	Stoneyhurst	5	19/10/2005	2	R	*AH*	1.94	1.4	71.2	19.1	81	4.2	33	21.8	1.50	92.1
2005	337	Stoneyhurst	379/02	19/10/2005	1	R	*AH*	2.76	2.0	75.3	19.2	79	4.9	36	25.3	1.58	85.4
2005	338	Stoneyhurst	32	19/10/2005	1	R	*BH*	3.00	2.0	67.0	20.7	65	3.2	45	15.5	1.00	99.5
2005	340	MV144	61	20/10/2005	1	E	*AB*	2.64	2.0	75.2	14.6	82	3.5	24	24.3	0.83	78.7
2005	342	MV144	143	20/10/2005	1	R	*BF*	2.26	1.8	81.2	16.3	87	2.9	23	17.9	0.33	84.5
2005	343	Stoneyhurst	298/02	20/10/2005	1	R	*AB*	2.26	1.6	70.6	18.4	83	3.7	11	19.9	0.50	95.1
2005	345	Stoneyhurst	365/02	20/10/2005	2	R	*GH*	2.44	2.0	75.1	19.3	71	5.3	21	27.3	4.08	79.6
2005	348	Stoneyhurst	355/02	20/10/2005	1	R	*AB*	2.58	1.8	73.5	18.2	77	3.3	36	18.2	0.33	99.1
2005	349	MV144	76	21/10/2005	1	E	*AG*	2.68	2.0	77.9	18.7	68	4.1	30	21.9	1.33	89.7
2005	350	Stoneyhurst	31	20/10/2005	1	R	*HH*	2.78	1.8	65.3	20.2	70	4.4	35	21.8	3.08	83.8
2005	351	MV144	142	22/10/2005	1	R	*AG*	2.26	1.8	79.6	15.5	73	3.2	45	20.5	0.42	82.9
2005	352	Stoneyhurst	383/02	22/10/2005	1	.	*GH*	2.74	2.0	75.3	18.0	83	3.7	17	20.6	0.83	85.8
2005	353	MV144	118	22/10/2005	2	E	*FG*	1.68	1.1	71.4	16.8	75	3.3	37	19.5	0.67	77.8
2005	354	MV144	118	22/10/2005	2	E	*FG*	2.66	2.0	75.5	16.0	87	3.1	20	19.2	0.33	94.5
2005	355	Stoneyhurst	21	23/10/2005	1	R	*AB*	2.24	1.8	80.2	17.5	77	3.1	33	17.5	0.17	84.2
2005	356	Stoneyhurst	361/02	25/10/2005	2	R	*BH*	2.30	1.6	71.6	18.2	74	4.2	36	23.2	1.67	95.7
2005	357	Stoneyhurst	361/02	25/10/2005	2	E	*AB*	2.20	1.4	64.7	17.7	77	4.7	18	26.4	2.33	97.3
2005	358	Stoneyhurst	4	25/10/2005	1	E	*AB*	2.28	1.5	70.2	20.4	70	4.1	39	20.2	2.17	101.4
2005	359	MV144	93	25/10/2005	1	.	*GH*	2.12	1.6	78.3	17.5	68	2.8	31	15.8	0.33	75.7
2005	360	MV144	130	26/10/2005	1	E	*GG*	2.34	1.8	79.2	18.0	86	2.8	26	15.8	0.33	78.0

2005	363	Stoneyhurst	383/02	28/10/2005	2	E	*AH*	2.08	1.6	77.8	17.0	84	3.2	42	18.6	0.50	89.0
2005	366	Stoneyhurst	8	4/11/2005	1	R	*AB*	2.76	2.3	83.4	18.3	85	3.6	52	19.7	1.17	75.4
2005	367	Stoneyhurst	144	2/11/2005	1	E	*AH*	2.40	1.6	65.9	19.4	82	4.7	16	24.2	2.75	89.8
2009	513	YT Southdown	822	.	2	R	*FH*	2.1	1.5	71.9	19.5	80	4.5	29	23.1	2.55	93.3
2009	514	sire 2	590	.	2	R	*BF*	2.6	2	76.6	25.1	90	6.2	15	24.7	20.5	73.8
2009	516	sire 2	509	.	2	R	*BF*	2.4	1.5	62.2	23.5	88	5.7	16	24.1	11.5	84.8
2009	518	sire 2	555	.	1	R	*AF*	2	1.5	76.2	20.1	73	4.1	22	20.4	1.1	79.9
2009	521	sire 2	591	.	1	R	*AF*	2.6	1.9	74.3	21.7	76	5.6	16	26	8.8	91.8
2009	523	sire 2	592	.	2	R	*FF*	2.2	1.6	70.8	19.6	93	4.6	20	23.5	3.25	101.6
2009	524	sire 2	616	.	1	R	*BF*	1.8	1.2	66.6	17.7	78	4.6	13	26.2	2.05	101.2
2009	525	sire 2	564	.	2	R	*AF*	1.8	1.3	74.8	18.3	66	3.7	14	20.5	0.8	98.2
2009	528	YT Southdown	427	.	1	R	*AF*	2.4	1.7	69.6	21.8	82	4.8	14	22	5.4	96.4
2009	529	YT Southdown	355	.	1	R	*FF*	1.7	1	61.1	19.6	71	4.3	12	21.9	1.55	129.7
2009	530	YT Southdown	822	.	2	R	*FF*	2.1	1.4	66.8	18.5	91	4.2	16	22.5	1.2	106.6
2009	532	YT Southdown	1505	.	2	R	*AF*	2.1	1.4	67.6	17	93	3.3	28	19.2	0.35	93.4
2009	533	YT Southdown	1096	.	1	R	*BF*	2.1	1.4	66.1	20.3	62	4.3	39	21.1	1.4	109.4
2009	534	YT Southdown	2603	.	1	R	*BF*	2.3	1.7	75.9	18	92	3.6	26	19.9	0.3	89.9
2009	535	YT Southdown	2026	.	2	R	*AF*	2.2	1.6	70.9	17.5	100	3.7	18	21.4	0.45	78.3
2009	536	YT Southdown	2610	.	2	R	*FF*	2.4	1.6	68.6	20	83	4.9	24	24.7	2.95	88.5
2009	537	YT Southdown	218	.	1	R	*FG*	2.1	1.5	69.6	20.9	75	4.9	18	23.5	3.9	106.8
2009	539	YT Southdown	222	.	1	R	*FG*	2.5	1.9	74.4	20.8	80	4.5	22	21.9	3.35	82
2009	540	YT Southdown	1488	.	1	R	*FH*	2.1	1.4	67.9	20.1	74	3.8	24	18.9	0.85	112.9
2009	541	YT Southdown	63	.	2	R	*AF*	2.6	1.7	63.8	21.8	79	5.2	27	23.9	5.2	71.3
2009	544	YT Southdown	6	.	2	R	*AF*	2.9	2.3	78.7	18.3	94	4	28	21.7	0.75	79.9
2009	546	YT Southdown	554	.	1	R	*FG*	2	1.5	75.4	17.6	79	4.1	28	23.3	1.1	104.9
2009	547	YT Southdown	1137	.	2	R	*AF*	2	1.4	68.8	20.2	78	5.3	18	26.1	5.1	97.4
2009	548	YT Southdown	2206	.	2	R	*AF*	2.1	1.4	66.6	20.8	72	4.1	35	19.9	1.5	95.9
2009	549	YT Southdown	448	.	2	R	*FG*	2	1.3	64.7	18	65	3.4	25	18.9	0.25	106.3
2009	550	YT Southdown	2009	.	2	R	*EF*	1.8	1.2	69.4	18.5	68	4.3	21	23.3	1.6	125.6
2009	551	sire 2	1182/04	.	1	R	*AF*	2.6	1.8	67.9	22.4	89	4.9	16	22	6.8	79.8
2009	552	sire 2	519	.	2	R	*FF*	2.2	1.6	73.2	18.5	89	4.3	27	23.3	1.1	68.7
2009	553	sire 2	454	.	1	R	*FG*	2.6	1.8	71	18	79	3.8	23	21.1	0.7	84.4
2009	554	sire 2	1191/04	.	1	R	*FH*	1.9	1.2	65.8	22	71	4.5	12	20.3	4.15	99
2009	555	sire 2	573	.	2	R	*FH*	2.5	1.7	69.5	21.9	87	5.2	17	23.6	7.1	93.6
2009	556	sire 2	531	.	2	R	*FG*	2.6	1.6	63.1	19.9	96	4.8	12	24	2.4	111.1
2009	557	sire 2	564	.	2	R	*BF*	2.2	1.6	72.1	19.6	71	4.7	10	24.1	2.75	112.1
2009	558	sire 2	531	.	2	R	*FF*	2.2	1.4	62.6	20.6	80	4.2	18	20.4	2.55	127.9
2009	559	sire 2	1192/04	.	2	R	*AF*	2.1	1.6	74.4	22.8	94	5.5	21	24.4	10.25	87.1
2009	560		1185/04	.	2	R	*BC*	1.3	1	75.9	20	83	5.6	14	28.2	5.6	98.4
2009	561		1185/04	.	2	R	*BC*	1.5	1.1	76.6	18.9	67	4.8	11	25.4	2	91.5

2009	562	sire 2	1192/04	.	2	R	*FF*	1.8	1.1	62.9	16.7	68	3.2	24	19	0.3	107.7
2009	563	sire 2	650	.	2	R	*BF*	1.8	1.3	72.4	17	85	4	24	23.6	0.1	86.3
2009	564	sire 2	261	.	1	R	*AF*	2.5	1.6	64.6	21	72	4.4	17	20.9	2.15	90.2
2009	565	sire 2	462	.	1	R	*FH*	1.5	1.1	72.6	20.9	68	4.7	23	22.5	3.6	95.2
2009	566	sire 2	266	.	1	R	*AF*	2.4	1.6	67.4	19.6	70	4.1	11	20.9	1.05	85.4
2009	567	sire 2	313	.	1	R	*FG*	1.8	1.2	68.2	18.5	76	4.1	14	21.9	1.25	107.8
2009	568	sire 2	268	.	2	R	*FF*	2.1	1.6	77.7	20.1	75	4.6	37	22.7	3	82.4
2009	569	sire 2	357/02	.	1	R	*BF*	2	1.4	69.7	18.9	57	4.4	34	23.5	1.65	90
2009	570	sire 2	489	.	1	R	*FF*	2.2	1.5	67	19.9	73	4.4	7	22.2	1.7	106.6
2009	571	sire 2	1065	.	2	R	*FH*	2.6	1.8	69.7	20.6	86	5.2	14	25.3	4.15	102.9
2009	572	sire 2	369/02	.	1	R	*AF*	2.5	1.9	75.2	18.9	85	4.4	27	23.2	1.15	82.8
2009	573	sire 2	480	.	1	R	*FF*	2.2	1.5	68.2	20.2	76	4.7	9	23.4	2.45	96.4
2009	574	sire 2	263	.	1	R	*FH*	2	1.5	72.7	19.5	77	4.6	14	23.8	2.2	95.2
2009	577	YT Southdown	2385	.	2	R	*BF*	2	1.3	66.1	20.5	75	5	14	24.3	2.85	139.7
2009	578	YT Southdown	586	.	2	R	*FF*	2.1	1.5	70.5	21.6	52	5.3	16	24.7	6.5	102.9
2009	579		6/07	.	1	R	*AC*	1	0.7	71.7	20.9	49	4.8	21	22.8	3.4	139.7
2009	582	YT Southdown	1449	.	1	R	*FH*	2.4	2	82.9	18.5	88	3.8	26	20.5	0.6	112.1
2009	583	YT Southdown	69	.	1	R	*FG*	2.6	1.5	59.2	20.9	61	5	14	23.9	5.35	106.8
2009	584		422	.	1	R	*AC*	1.8	1.3	73.5	18.4	84	4.1	22	22.1	1	72.6
2009	585	YT Southdown	850	.	2	R	*FH*	2.1	1.3	60.1	20.1	65	4.5	12	22.2	1.75	114.5
2009	586	YT Southdown	435	.	2	R	*FH*	2.4	1.6	68.2	20	73	4.1	21	20.5	0.6	118.8
2009	587	YT Southdown	2591	.	2	R	*FF*	2.1	1.4	69	20.3	78	4.8	15	23.5	2.55	117.3
2009	588	YT Southdown	2591	.	2	R	*FF*	2.3	1.4	59.7	22.2	69	5	24	22.5	5.45	146.7
2009	589	YT Southdown	14/07	.	1	R	*CF*	1.6	1.1	69.2	19.9	72	4.5	15	22.8	1.4	82.8
2009	591	YT Southdown	1341	.	1	R	*FH*	1.4	1.1	76.6	17.5	74	3.3	26	19.1	0.15	107.3
2009	592	YT Southdown	1609	.	2	R	*FH*	2.4	1.4	57.7	22.3	63	4.5	29	20.1	4.15	78.6
2009	594	YT Southdown	47/07	.	1	R	*CF*	2	1.4	70.9	23.5	78	7	25	29.6	17.6	102.5
2009	597	YT Southdown	1281	.	1	R	*AF*	2.3	1.5	66.7	21.5	74	4.7	23	21.6	3.3	129.6
2009	598	YT Southdown	2095	.	1	R	*AF*	2	1.3	63.9	19.9	83	4.3	15	21.8	1.25	109.3
2009	602	sire 2	1186/04	.	1	R	*FG*	2.1	1.4	65.8	21.2	60	4.9	17	23.2	3.85	108.3
2009	603	sire 2	514	.	2	R	*FH*	2	1.4	71.6	18	78	4.5	21	25.3	1	80
2009	604	sire 2	280	.	1	R	*FH*	1.9	1.3	68.2	20.2	59	4.6	22	23	2.9	97.1
2009	606	sire 2	594	.	1	R	*AF*	2.2	1.6	74.4	21.8	69	4.8	17	21.9	4.1	121.2
2009	607	sire 2	331	.	1	R	*AF*	2.7	1.9	71.8	16.8	79	3.8	22	22.4	0.6	81.3
2009	608	sire 2	335	.	1	R	*BF*	2.5	1.9	74.1	18.1	84	4.4	17	24.5	1	100.7
2009	609	sire 2	279	.	1	R	*FF*	1.8	1.3	72.7	18.8	68	4.5	24	24.1	1.25	80.6
2009	610	sire 2	282	.	2	R	*FG*	1.9	1.4	75.8	18.1	82	4.2	19	23	0.55	73.6
2009	611	sire 2	282	.	2	R	*FG*	1.8	1.2	64.3	17.7	78	3.8	14	21.6	0.6	111.1
2009	612	sire 2	492	.	1	R	*FH*	1.9	1.5	77	17.1	73	4.1	21	23.8	0.7	100
2009	613	sire 2	483	.	1	R	*FH*	2.1	1.3	63.5	20.2	68	4.5	21	22	2.3	93.2

2009	614	sire 2	402	.	1	R	*FG*	1.8	1.4	75.1	18.4	70	4.4	28	24	1.65	84.4
2009	770	YT Southdown	256	.	1	E	*BF*	2.6	2	77.3	18.5	112	3.4	25	18.4	0.25	74.9
2009	771	YT Southdown	1267	.	1	E	*FG*	2.1	1.3	63.1	22	89	4.4	28	19.9	2.75	116.3
2009	776	YT Southdown	2622	.	1	E	*FH*	2.2	1.5	67.6	21	72	4.5	25	21.5	2.95	89.8
2009	777	YT Southdown	1903	.	1	E	*BF*	2.3	1.7	74.4	20.8	88	4.8	28	22.8	3.95	95.3
2009	778	YT Southdown	2168	.	1	E	*FG*	2	1.6	80	17.5	73	4.6	22	26.1	1.2	95.3
2009	779	YT Southdown	6	.	2	E	*FF*	2.7	1.7	61.2	19.2	82	4.9	23	25.7	1.25	93.4
2009	780	YT Southdown	1684	.	1	E	*FG*	2.5	1.7	66.4	20.7	76	5.5	15	26.4	4.9	89.5
2009	781	YT Southdown	2206	.	2	E	*AF*	2.1	1.4	65.8	22.1	70	5.6	30	25.2	8	92.9
2009	783	YT Southdown	713	.	1	E	*AF*	2	1.5	73.1	19.4	75	4.5	17	23.4	0.95	88.8
2009	784	YT Southdown	1137	.	2	E	*FG*	2	1.3	64.9	17.7	93	4	24	22.5	0.4	100.9
2009	785	YT Southdown	2343	.	1	E	*AF*	2.6	1.8	69	19.6	93	4.3	25	21.9	1.5	109.1
2009	786	YT Southdown	1505	.	2	E	*AF*	2.1	1.4	64.3	20.4	72	4.7	38	23.2	3.25	99
2009	787	YT Southdown	157	.	1	E	*FF*	2	1.2	61.9	22.2	77	4.6	15	20.9	4.2	105.1
2009	788	YT Southdown	2026	.	2	E	*FH*	2.4	1.4	58	20.4	72	5.1	16	25	4.85	119
2009	789	YT Southdown	2610	.	2	E	*FF*	1.8	1.2	68.5	19.2	61	4.3	14	22.4	1.85	117.2
2009	790	YT Southdown	611	.	1	E	*CF*	1.5	1.2	82.4	17.9	74	4	20	22.6	0.35	82.3
2009	791	YT Southdown	121	.	1	E	*FF*	2.1	1.2	56.3	20.5	64	4.8	16	23.3	3.1	111.2
2009	801	YT Southdown	740	.	1	E	*FH*	2.4	1.7	71.4	20.8	96	5.1	20	24.6	5.75	93.8
2009	802	YT Southdown	2009	.	2	E	*FG*	1.8	1.2	64.4	20.6	76	4.1	26	19.7	1.85	104.9
2009	804	YT Southdown	448	.	2	E	*FH*	1.7	1	59	18.5	84	3.6	22	19.7	0.6	113
2009	805	YT Southdown	1266	.	1	E	*FG*	2.3	1.7	72.8	19.9	88	4.3	24	21.4	1.95	79.7
2009	807	YT Southdown	586	.	2	E	*FG*	2.6	1.8	70.4	19.7	97	4.3	17	21.8	1.4	83.5
2009	808	YT Southdown	1696	.	1	E	*FG*	2.1	1.4	68.5	21.5	68	4.5	16	20.7	2.85	134.5
2009	813		417	.	1	E	*CG*	1.6	1.2	75.7	19.9	76	4.6	26	23.1	1.6	106.4
2009	814	YT Southdown	522	.	2	E	*FG*	1.9	1.2	63.9	19.3	71	4.2	39	21.7	1.7	102
2009	815	YT Southdown	453	.	1	E	*FF*	2.3	1.6	67.7	20.6	89	4.4	27	21.5	2.3	83.5
2009	818	YT Southdown	1389	.	2	E	*AF*	2	1.4	70.2	22.7	75	5.2	35	22.8	9	90.8
2009	819	YT Southdown	850	.	2	E	*BF*	2.4	1.6	66.8	20.3	90	5	17	24.8	2.4	99.9
2009	820	YT Southdown	540	.	1	E	*BF*	2.2	1.5	68.4	19.7	78	4.8	22	24.6	2.55	106.3
2009	821	YT Southdown	2498	.	1	E	*BF*	2.2	1.5	70.3	20.8	63	4.1	31	19.6	1.85	108
2009	822	YT Southdown	435	.	2	E	*FH*	2.4	1.9	79.1	20.4	84	4.7	35	22.9	1.6	79.4
2009	823	YT Southdown	373 Pink	.	1	E	*BF*	2.6	1.6	63	21.1	93	4.5	19	21.2	3.25	115.6
2009	826	YT Southdown	2147	.	2	E	*FG*	2.1	1.4	67.1	19	74	4.8	12	25.5	2.45	102.3
2009	827	YT Southdown	1079	.	1	E	*AF*	2.6	1.8	70.7	21.4	78	5	24	23.2	4.7	103.9
2009	828	YT Southdown	1609	.	2	E	*FG*	1.7	1.1	64.1	18.9	90	4.5	15	24	1	110
2009	830	YT Southdown	28/07	.	1	E	*CF*	1	0.7	70.6	21.1	49	5.1	32	24.4	4.05	94.1
2009	832	YT Southdown	1901	.	1	E	*FF*	2.6	1.9	73.1	20.3	89	4.5	30	22.3	1.4	71.7
2009	833	YT Southdown	182	.	2	E	*AF*	2.7	2.1	77	19.9	104	4.5	22	22.8	2.1	87.1
2009	835	YT Southdown	525	.	2	E	*BF*	2.5	1.8	73.6	19.9	96	3.9	24	19.6	0.9	98.4

2009	836	YT Southdown	525	.	2	E	*BF*	2.6	1.9	72.8	20.8	91	4.7	16	22.6	3.15	92
2009	837	YT Southdown	316	.	1	E	*FJ*	2	1.2	61.7	19.5	76	4.1	30	21.2	1.05	144.2
2009	838	YT Southdown	255	.	2	E	*FG*	1.8	1.2	64.1	19.6	77	3.9	18	20.1	0.4	127.9
2009	839	YT Southdown	255	.	2	E	*AF*	2.1	1.5	71.3	19.8	86	4.7	24	24	2.8	88.5
2009	841	YT Southdown	109	.	2	E	*FF*	2.1	1.4	68.1	18.7	86	4.6	8	24.9	1.5	89.2
2009	842	YT Southdown	109	.	2	E	*FF*	2.3	1.5	66.6	20.3	76	4.6	15	22.9	2.35	82.9
2009	843	sire 2	1055	.	2	E	*AF*	2.8	2.1	73.9	18.7	80	4.4	18	23.5	1.2	86
2009	844	sire 2	1055	.	2	E	*AF*	2.6	1.8	67.3	22.7	82	6.1	18	26.8	11.25	92.5
2009	845	sire 2	575	.	1	E	*FH*	2.4	1.7	68.8	19.8	103	4	19	20.4	1.05	87.9
2009	847	sire 2	569	.	1	E	*FH*	2.3	1.6	68.3	22.9	63	5.4	22	23.6	7.55	90
2009	850	sire 2	1029	.	1	E	*FH*	2.5	1.8	70.6	21.8	86	5.3	34	24.1	7.25	90.1
2009	852	sire 2	597	.	1	E	*CF*	1.4	1	71.9	20.8	88	5.4	12	25.7	4.8	91.8
2009	853	sire 2	592	.	2	E	*FG*	2.2	1.4	65	23.6	91	6.3	18	26.7	14.6	95.1
2009	854	sire 2	571	.	2	E	*AF*	2.9	2.1	73.3	18.9	78	3.6	23	19	0.35	112.3
2009	855	sire 2	571	.	2	E	*AF*	2.1	1.2	56.7	21.8	60	5.2	14	23.8	7.4	100
2009	857	sire 2	548	.	1	E	*CF*	1.5	1.1	72.4	21.5	61	5.5	20	25.5	6.6	112
2009	858	sire 2	460	.	1	E	*FH*	1.8	1.3	70	19	83	3.9	28	20.6	0.7	97.7
2009	859	sire 2	650	.	2	E	*BF*	1.6	1.3	83	18	72	5.1	24	28.2	1.1	84.2
2009	860	sire 2	384/02	.	1	E	*FG*	2.2	1.4	65.5	21.9	78	5	16	22.9	6.8	91.9
2009	861	sire 2	464	.	1	E	*AF*	2.2	1.5	68.5	21.3	75	5.3	17	25	5.75	106.6
2009	862	sire 2	596	.	1	E	*AF*	2.5	1.8	72.4	20.7	80	4.1	18	20	2.3	76.9
2009	863	sire 2	405	.	1	E	*FG*	2.6	1.8	68.2	19.1	77	4.4	19	23.1	1.5	102.2
2009	865	sire 2	573	.	2	E	*FH*	2.1	1.5	70.9	20.3	85	4.7	21	23	2.8	91.2
2009	866	sire 2	268	.	2	E	*FF*	2.3	1.8	77.3	19.6	82	4.9	25	25	3.25	92.8
2009	867	sire 2	563	.	1	E	*FG*	2	1.4	69.5	21.8	74	5.9	18	26.9	9.3	99.1
2009	868	sire 2	286	.	1	E	*BF*	2	1.4	70.1	19.7	70	4.7	21	24.1	2.2	84.9
2009	869	sire 2	272	.	1	E	*BF*	2.2	1.6	72.7	20.5	76	4.7	28	22.7	3.15	88
2009	870	sire 2	1065	.	2	E	*FH*	1.9	1.2	65.3	20	79	4.9	35	24.5	3.75	98.7
2009	871	sire 2	601	.	1	E	*BF*	2.4	1.8	77.3	19.3	85	4.2	29	21.5	0.85	77.8
2009	874	sire 2	2141	.	1	E	*BF*	2.3	1.4	59.8	20.9	75	5.2	22	24.7	4.15	127.7
2009	876	sire 2	404	.	1	E	*FJ*	2.3	1.7	74.2	18.2	88	5	24	27.6	2.6	76.8
2009	877	sire 2	514	.	2	E	*BF*	2.2	1.7	75.9	19.3	89	4.3	21	22.2	1.55	76.8
2009	878	sire 2	515	.	1	E	*AF*	2.5	2	81.9	21.9	86	5.3	40	23.9	5.4	80.1
2009	879	sire 2	275	.	2	E	*FF*	1.8	1.1	63.3	17.1	83	3.8	20	22.4	0.75	130.1
2009	880	sire 2	458	.	1	E	*FF*	2	1.3	67.1	22.5	63	4.8	22	21.3	6	91.6
2009	881		379/02	.	1	E	*HH*	1.5	1.1	76.2	19.1	70	4.3	13	22.8	1.7	100.2
2009	884	sire 2	468	.	1	E	*FH*	1.9	1.3	69.5	18.5	74	4.4	24	23.7	1.8	98.6
2009	885	sire 2	265	.	1	E	*AF*	2.2	1.6	71.3	20.2	78	4.6	20	22.6	2.15	96.4
2009	886	sire 2	576	.	1	E	*FH*	2	1.4	70.7	16.8	84	4.8	18	28.6	1	76.5
2010	752	Merino	863		1	.	*AF*	2.86	2.13	74.5	17.3	90	3.6	22	21.1	0.25	67.3

2010	753	Merino	47/07	2	.	AF	3.16	2.68	84.7	19.3	91	3.8	25	19.9	1.25	85.5
2010	754	Merino	47/07	2	E	AG	3.22	2.24	69.7	22	94	4.2	20	19.2	3.95	94
2010	756	Merino	788	1	.	AF	2.66	2.19	82.5	17.2	97	3.7	26	21.8	0.6	67.3
2010	757	Merino	758	1	E	AF	2.78	2.38	85.7	18.2	88	4.1	21	22.5	1.65	68.2
2010	759	Merino	734	1	E	AF	3.34	2.57	76.8	19.8	82	3.6	38	18.2	0.9	79.2
2010	760	Merino	816	1	.	AF	2.54	2.12	83.6	18.6	75	3.5	24	19	0.9	99.4
2010	761	Merino	877	1	E	AF	1.78	1.44	80.7	17.5	77	3.6	30	20.4	0.55	86.5
2010	765	Merino	519	1	E	AA	3.16	2.25	71.2	16.1	86	2.8	21	17.7	0.2	75.9
2010	766	Merino	94/07	1	E	AF	2.72	1.65	60.8	25.7	71	6	18	23.3	20.4	99.4
2010	767	Merino	548	1	E	FF	3.52	2.38	67.7	25.2	67	5	39	19.8	13.3	80.5
2010	768	Merino	837	1	.	AB	2.54	2.08	82	18.3	95	3.9	26	21.1	1	80.5
2010	769	Merino	745	2	.	AF	2.48	1.97	79.5	18.1	90	4.6	36	25.6	1.8	76.4
2010	770	Merino	745	2	E	AF	2.22	1.72	77.7	18.4	63	3.6	34	19.8	0.45	90.1
2010	771	Merino	723	1	E	AA	2.14	1.32	61.9	18.5	66	3.9	15	21	1.1	110.9
2010	772	Merino	569	1	E	AF	2.96	2.25	75.9	24	69	5.2	24	21.7	8.6	104
2010	775	Merino	95	1	E	AF	3.12	2.53	81	20.2	87	4.5	17	22.3	2.6	94.4
2010	776	Merino	42	2	E	AF	2.66	2.02	76.1	17.5	98	3.1	34	17.6	0.15	76.7
2010	777	Merino	42	2	.	AF	2.8	1.84	65.6	19	86	3.1	28	16.2	0.35	82.6
2010	778	Merino	24	1	E	AF	3.48	2.85	81.9	17.3	80	3.9	20	22.5	1.35	66.5
2010	779	Merino	84	2	E	AF	3.06	2.59	84.7	19.3	77	4.4	27	22.7	2.4	75.2
2010	780	Merino	84	2	.	AG	2.76	2.00	72.4	18.2	78	3.5	20	19.5	0.7	73.5
2010	781	Merino	740	1	E	AF	2.66	1.91	71.7	19.5	99	3.9	37	20	1.5	66.5
2010	782	Merino	835	1	E	AA	2.92	2.23	76.3	19.3	70	4.5	16	23.4	2.4	87.8
2010	783	Merino	753	1	.	AF	2.26	1.59	70.5	18.8	96	3.9	24	20.9	0.75	96.1
2010	784	Merino	796	1	E	AF	2.62	1.80	68.6	19.8	82	3.5	17	17.5	0.8	91.3
2010	786	Merino	781	1	E	AF	3.44	2.74	79.6	20.7	104	4.4	29	21.3	1.35	81.7
2010	787	Merino	854	1	.	AF	2.28	1.65	72.5	18.2	74	3.6	34	19.5	0.6	95.1
2010	788	Merino	789	1	.	AF	3.34	2.31	69.3	18.5	83	4.1	23	22.3	1.95	66.7
2010	790	Merino	50	1	.	AB	2.98	2.54	85.4	18.8	108	4.2	11	22.6	1.4	65.5
2010	791	Merino	36	2	.	AF	2.58	1.99	77.2	17.3	65	3.3	35	19	0.2	71.7
2010	792	Merino	36	2	.	AF	2.42	2.04	84.1	18.8	70	3.7	33	19.6	1.05	77
2010	794	Merino	742	2	.	AA	2.42	2.00	82.7	23	108	5.7	25	24.9	7.15	76
2010	795	Merino	857	2	E	AF	2.68	2.13	79.4	18.3	110	3.8	21	20.9	0.55	71.4
2010	796	Merino	857	2	E	AF	2.38	1.79	75.3	21.5	82	4.1	26	18.9	1.85	92.2
2010	797	Merino	880	1	E	AH	3.96	2.61	66	18.8	97	3.8	30	20.1	1	83.9
2010	798	Merino	824	1	.	AA	2.52	1.88	74.7	16.7	97	3.8	30	22.6	0.4	73.6
2010	799	Merino	650	1	.	AB	2.54	1.93	75.8	16.4	98	3.8	22	22.9	0.55	72.5
2010	801	Merino	70	2	.	AF	2.82	2.14	75.8	18.1	109	3.8	25	20.9	0.55	71.6
2010	802	Merino	70	2	E	AF	2.84	2.39	84	18.6	99	3.5	18	18.8	1	72.1
2010	804	Southdown	32	2	E	BF	3.28	2.21	67.5	22.4	88	4.1	25	18.5	3.1	83.7

2010	805	Southdown	32	2	.	*BF*	2.62	1.89	72.3	20.5	88	4.4	24	21.3	2.8	96.2
2010	806	Southdown	218	1	E	*FG*	2.66	1.96	73.8	22.6	75	4.5	25	20	5.15	109.2
2010	807	Southdown	1470	1	.	*FF*	2.62	1.94	73.9	20.7	99	3.9	26	18.7	1	87.3
2010	808	Southdown	727	1	.	*AF*	2.72	1.85	68	19.8	86	4.5	18	22.8	1.4	109.6
2010	809	Southdown	1049	1	E	*FG*	2.46	1.47	59.9	21.3	75	4	22	18.6	1.6	87.7
2010	810	Southdown	24	1	.	*FG*	2.88	1.85	64.2	20.4	89	4.4	26	21.7	3.3	91.9
2010	811	Southdown	166	1	E	*AF*	2.32	1.54	66.2	21.4	84	3.8	22	17.8	1.8	100
2010	812	Southdown	525	1	.	*AF*	3.34	2.33	69.8	20	88	4.5	29	22.3	2.6	72.5
2010	813	Merino	68	1	.	*AF*	2.68	1.74	65.1	18	87	3.7	24	20.7	0.85	96
2010	814	Southdown	531	2	E	*AF*	2.38	1.50	62.9	23.7	71	5.3	20	22.3	10	87.3
2010	816	Southdown	525	1	.	*FG*	2.3	1.59	69	20.4	77	4.2	24	20.8	1.95	98
2010	821	Southdown	564	1	.	*AF*	2.52	1.80	71.6	21.3	57	5.1	25	24.1	4.8	102
2010	822	Southdown	1186/04	2	.	*FH*	2.36	1.65	70	25.3	76	5.3	34	21	16.1	87.3
2010	825	Southdown	565	1	.	*FG*	1.86	1.38	74.3	20.2	74	4.6	24	23	1.4	94.2
2010	827	Southdown	1294	2	.	*FF*	2.14	1.49	69.8	20.8	75	5.1	21	24.7	4.75	77.9
2010	828	Southdown	563	1	E	*AG*	2.94	2.22	75.5	18.3	105	3.8	27	20.5	0.85	68.8
2010	830	Southdown	1182/04	1	.	*FH*	2.32	1.82	78.3	25.3	96	6.3	12	24.8	21.75	76.7
2010	831	Southdown	843	1	E	*FF*	3.04	2.31	75.9	20.2	105	4.1	31	20.4	1.45	89.2
2010	833	Southdown	101	2	E	*AF*	2.64	2.23	84.5	21.8	70	5.4	22	25	4.7	96.3
2010	836	Southdown	23	1	.	*AG*	2.42	1.74	72.1	16.5	102	3.8	19	23	0.25	79.1
2010	839	Southdown	76	2	.	*BF*	1.94	1.54	79.6	20.7	86	4.3	18	20.8	2.8	99.6
2010	840	Southdown	198	2	.	*BF*	2.68	1.90	70.8	20.6	100	4.1	36	19.7	1.2	81.7
2010	841	Southdown	198	2	.	*FG*	2.46	1.55	63.1	21.1	84	4.7	19	22.5	3.95	99.5
2010	842	Southdown	275	1	.	*FF*	2.66	1.76	66.2	19.5	85	4.3	21	22.1	1.4	83.1
2010	846	Southdown	85	2	E	*AF*	3.04	2.20	72.3	25.7	84	6.2	21	24.2	19.95	85.1
2010	847	Southdown	85	2	E	*FG*	3.02	2.12	70.3	24.9	89	4.4	35	17.5	9.1	91.5
2010	850	Southdown	169	1	E	*FH*	2.3	1.47	63.9	21.6	68	4	34	18.7	2.45	95.7
2010	851	Merino	14	1	.	*AF*	3.4	2.62	77	19.7	90	4.9	25	25.1	2	81.2
2010	852	Merino	871	1	E	*AF*	2.64	2.11	79.8	18.6	84	3.2	35	17.2	0.6	63.4
2010	853	Merino	792	1	.	*AF*	2.08	1.65	79.3	19.5	71	4.5	32	22.9	2.75	100.9
2010	856	Merino	29	2	E	*AF*	2.84	2.14	75.3	18.7	80	3.5	22	18.6	0.2	76.3
2010	857	Merino	29	2	E	*AF*	2.88	2.17	75.2	18.2	85	3.2	25	17.8	0.3	76.7
2010	858	Merino	776	1	E	*AB*	2.88	2.34	81.2	17.8	100	2.9	31	16.1	0.3	73.6
2010	860	Merino	802	1	.	*AF*	2.96	2.52	85.1	18.9	92	4.2	14	22.2	2.35	77.9
2010	861	Merino	805	2	.	*AF*	2.54	2.07	81.3	19.1	106	3.5	41	18.2	0.85	72
2010	862	Merino	805	2	E	*AF*	2.78	2.40	86.5	19.3	107	4.8	33	24.9	2.75	75.3
2010	863	Merino	830	1	E	*AB*	3.6	2.93	81.4	18.9	86	3.3	40	17.4	0.4	70.3
2010	864	Merino	836	2	E	*AA*	2.48	1.98	80	18.7	85	3.3	36	17.8	0.25	76.8
2010	865	Merino	836	2	.	*AF*	2.96	2.44	82.4	19.5	76	4.7	18	24.2	2.35	69.5
2010	866	Merino	35	1	.	*AF*	2.8	1.84	65.6	22.1	87	4	18	18.2	2.25	117.6

2010	868	Merino	766	1	E	*AF*	3.3	2.29	69.5	20.2	78	3.9	29	19.4	1.85	83.1
2010	871	Merino	1198/04	1	.	*FG*	2.3	1.61	70	20.6	71	4.5	28	21.6	1.15	83.5
2010	872	Merino	6	2	.	*AF*	2.58	2.04	79	18.7	108	4	22	21.2	0.7	68.8
2010	873	Merino	6	2	.	*AF*	2.72	1.91	70.4	18.4	80	3.3	18	18.1	0.2	77.5
2010	874	Merino	901	1	E	*AA*	3.62	2.71	74.8	20	104	4	23	20	1.4	77
2010	877	Merino	758	1	E	*AA*	2.22	1.79	80.5	19	91	4.1	27	21.4	1.2	97.3
2010	878	Merino	741	1	.	*AF*	2.44	2.03	83.2	20.4	89	4.4	27	21.6	2.65	78.8
2010	879	Merino	73	1	.	*AA*	3.3	2.32	70.3	19.2	99	4.4	16	22.6	1.5	82.8
2010	881	Merino	12	1	.	*AF*	2.5	2.08	83.2	16.2	84	3.1	17	19.3	0.25	64.2
2010	885	Merino	1048	1	.	*BF*	2.28	1.71	75	19.9	83	4.2	20	21.3	1.65	89
2010	886	Merino	26	1	E	*AG*	2.9	2.29	78.9	18.5	100	3.3	15	18	0.2	76.9
2010	887	Merino	732	1	E	*AF*	3.48	2.54	73	20.6	84	4.5	20	22	3.35	90.7
2010	889	Merino	872	1	.	*AG*	2.36	1.92	81.5	18.4	85	3.8	18	20.9	0.65	76.9
2010	890	Merino	10	2	.	*AF*	3.12	2.06	66.1	18	103	3.9	20	21.5	0.2	74.1
2010	891	Merino	10	2	.	*AF*	2.78	2.25	81	19.3	92	4.3	22	22.2	1.9	91.4
2010	892	Merino	878	1	.	*AF*	2.72	2.08	76.4	16.9	94	3.9	25	23.3	0.6	75.6
2010	895	Merino	882	1	.	*AA*	2.5	2.00	79.8	18.2	93	4.4	38	24.3	1.1	81.1
2010	896	Merino	28	1	.	*AF*	3.02	2.49	82.5	19.1	80	3.8	17	19.6	0.65	73.7
2010	897	Merino	814	1	.	*AA*	2.62	2.17	82.8	17.6	96	3.7	21	20.9	0.4	71
2010	898	Southdown	172	1	.	*AF*	2.2	1.65	75	20.3	76	4.2	24	20.8	1.8	100.3
2010	899	Southdown	2052	2	E	*BF*	3.16	2.28	72	20.4	84	4.4	29	21.7	3.1	90.5
2010	900	Southdown	2052	2	E	*AF*	2.68	1.75	65.3	21.8	64	5.1	27	23.4	6.8	104.1
2010	3151	Southdown	170	1	E	*FG*	2.82	1.93	68.3	24.6	74	4.6	29	18.7	10.9	98.2
2010	3154	Southdown	3153	2	.	*AF*	2.8	1.96	69.9	23.7	70	5.3	26	22.6	11.95	79.8
2010	3155	Southdown	3153	2	.	*BF*	2.38	1.60	67.4	18.8	80	3.7	21	19.6	0.4	107.9
2010	3156	Southdown	803	1	.	*BF*	2.14	1.49	69.8	20.5	73	4.1	38	19.9	1.7	89.4
2010	3158	Southdown	570	1	.	*AF*	2.58	1.71	66.4	18.9	84	3.9	16	20.9	0.95	106
2010	3159	Southdown	596	1	.	*AF*	2.4	1.86	77.5	14.9	74	2.8	27	19.1		64.1
2010	3161	Southdown	161	2	.	*BF*	2	1.28	63.8	19.6	94	3.5	17	17.8	0.8	96.2
2010	3162	Southdown	161	2	.	*FF*	2.14	1.26	59	21.3	77	4.4	31	20.8	4.35	114.8
2010	3163	Southdown	2591	1	.	*BF*	3.3	2.24	67.9	20.5	88	4.2	30	20.6	1.85	91.2
2010	3164	Southdown	1032	1	.	*FF*	2.5	1.74	69.5	20	86	4.9	20	24.3	3.45	95.3
2010	3166	Southdown	803	1	E	*BF*	2.56	2.07	81	23.2	81	5.3	27	22.7	9.5	105.1
2010	3174	Southdown	591	1	.	*AA*	3.34	2.32	69.4	23.7	76	5.8	17	24.4	12.1	103.7
2010	3175	Southdown	55	2	.	*AG*	3.3	2.82	85.5	20.7	81	5.1	26	24.4	4.55	85.4

2. *KRTAP6-1* genotype and wool data

Year	LambID	SireID	EweID	DOB	Birth rank	Sex	*KRTAP6-1*	GFW	CFW	Yield	MFD	MSL	FDSD	MSS	CVFD	PF	MFC
2005	202	MV144	105	1/10/2005	1	R	*BB*	3.00	2.2	73.4	17.2	86	3.9	24	22.5	1.00	88.2
2005	204	Stoneyhurst	370/02	1/10/2005	2	R	*AB*	2.34	1.6	70.6	19.0	93	3.8	26	20.3	1.17	81.5
2005	205	Stoneyhurst	389/02	2/10/2005	1	E	*BC*	2.82	1.9	66.5	20.9	80	5.0	18	23.7	5.25	82.9
2005	207	Stoneyhurst	9	2/10/2005	1	R	*BC*	3.62	2.5	68.3	19.3	60	3.6	40	18.5	1.67	85.6
2005	210	MV144	133	3/10/2005	2	R	*AB*	2.62	1.9	73.6	16.1	75	3.8	26	23.5	0.25	92.7
2005	216	MV144	85	4/10/2005	2	.	*BB*	2.86	2.3	83.7	17.2	91	5.6	23	32.6	2.58	88.8
2005	217	Stoneyhurst	16	4/10/2005	1	.	*AC*	2.50	1.8	73.4	20.7	76	4.7	45	22.7	4.50	87.8
2005	218	MV144	104	2/10/2005	1	R	*AB*	2.66	2.1	81.8	17.2	68	3.1	24	18.0	0.50	73.8
2005	219	MV144	97	28/09/2005	1	E	*AB*	2.22	1.6	74.5	18.3	76	4.0	21	21.6	1.33	101.0
2005	221	Stoneyhurst	150	4/10/2005	2	.	*BC*	3.32	2.2	68.1	24.7	94	7.1	14	28.8	22.58	90.3
2005	223	MV144	129	4/10/2005	2	E	*AB*	2.80	2.3	80.5	15.5	94	3.4	38	21.7	0.50	64.3
2005	224	MV144	110	4/10/2005	1	E	*AB*	2.82	2.1	76.3	17.4	75	3.5	54	20.1	0.83	88.3
2005	225	MV144	66	5/10/2005	1	R	*AB*	2.84	2.3	81.6	18.8	84	3.4	26	18.2	0.67	89.2
2005	227	Stoneyhurst	369/02	5/10/2005	2	R	*BC*	2.38	1.9	67.2	20.2	84	4.6	31	23.0	3.33	96.1
2005	228	Stoneyhurst	369/02	5/10/2005	2	R	*AB*	2.20	1.4	64.1	18.6	79	3.5	10	18.9	0.92	94.1
2005	229	MV144	64	4/10/2005	2	R	*AB*	2.04	1.5	76.0	16.8	92	3.1	26	18.2	0.50	87.5
2005	230	MV144	64	4/10/2005	2	E	*AB*	1.90	1.5	78.4	18.0	82	4.0	21	22.1	1.08	80.3
2005	231	MV144	121	7/10/2005	1	E	*AB*	2.50	2.0	78.2	16.3	81	2.8	32	17.2	0.00	93.8
2005	232	MV144	53	7/10/2005	1	R	*BB*	2.72	2.0	74.6	14.6	71	3.8	28	25.8	0.17	95.3
2005	234	Stoneyhurst	40	7/10/2005	1	E	*AC*	3.12	2.1	68.9	23.4	91	5.2	28	22.1	9.92	84.3
2005	235	Stoneyhurst	1	7/10/2005	1	E	*AB*	2.38	1.6	70.0	21.4	73	3.6	38	16.8	1.08	119.9
2005	236	MV144	134	7/10/2005	1	E	*AB*	2.04	1.5	77.4	17.0	88	3.2	18	18.8	0.17	86.6
2005	237	MV144	79	8/10/2005	1	R	*AB*	2.94	2.3	77.9	17.7	71	3.1	36	17.8	0.17	87.9
2005	238	MV144	80	8/10/2005	1	R	*AB*	2.88	2.3	81.3	15.7	72	2.9	54	18.5	0.08	92.1
2005	239	Stoneyhurst	145	7/10/2005	1	E	*AA*	2.48	1.5	63.2	19.8	69	3.8	17	19.3	0.92	82.1
2005	241	Stoneyhurst	380/02	7/10/2005	2	E	*AB*	2.10	1.5	70.3	17.1	61	3.9	32	22.7	0.42	105.2
2005	242	Stoneyhurst	392/02	7/10/2005	2	R	*BC*	2.50	1.7	67.7	21.3	80	5.4	20	25.1	5.17	96.8
2005	243	Stoneyhurst	392/02	7/10/2005	2	E	*BC*	2.30	1.6	68.0	21.3	82	4.3	31	20.0	4.00	87.7
2005	244	MV144	62	6/10/2005	1	R	*AB*	2.76	2.2	79.8	14.9	72	2.6	36	17.5	0.50	88.5
2005	245	Stoneyhurst	12	6/10/2005	2	R	*AA*	2.42	1.8	74.7	20.6	81	4.1	45	20.0	1.58	82.6
2005	247	Stoneyhurst	296/02	8/10/2005	2	E	*AB*	2.08	1.3	67.2	16.3	68	3.1	26	19.0	0.08	82.1
2005	248	Stoneyhurst	296/02	8/10/2005	2	E	*AC*	2.52	1.9	75.6	20.2	69	3.7	45	18.5	1.25	82.3
2005	250	MV144	72	18/10/2005	2	E	*AB*	2.38	1.7	74.3	17.6	80	4.0	33	23.0	1.08	88.0
2005	251	Stoneyhurst	297/02	9/10/2005	1	R	*BC*	2.76	2.0	72.4	19.3	90	2.7	39	13.9	0.00	89.5
2005	252	MV144	140	9/10/2005	1	E	*BC*	3.26	1.9	59.0	20.7	118	4.1	7	19.9	3.08	93.7
2005	253	MV144	99	9/10/2005	1	R	*BB*	2.98	2.3	79.0	16.3	83	3.3	33	19.9	0.17	78.9
2005	254	Stoneyhurst	358/02	9/10/2005	1	R	*AA*	2.84	2.1	76.0	18.2	82	3.1	29	17.1	0.17	99.4

2005	255	MV144	138	9/10/2005	1	E	*AB*	2.60	2.0	75.5	16.0	79	2.8	36	17.6	0.42	95.5
2005	256	MV144	141	9/10/2005	1	R	*AB*	2.74	1.9	68.8	16.4	80	3.3	34	20.3	0.50	98.4
2005	257	Stoneyhurst	376/02	9/10/2005	2	R	*BC*	2.26	1.5	67.9	20.0	78	5.7	13	28.3	6.33	101.8
2005	258	Stoneyhurst	376/02	9/10/2005	2	E	*AC*	2.48	1.9	77.4	18.0	76	4.2	16	23.0	1.08	92.6
2005	259	Stoneyhurst	2	9/10/2005	1	E	*AA*	2.86	2.1	74.5	20.4	96	5.5	32	26.7	7.17	79.2
2005	260	Stoneyhurst	373/02	10/10/2005	1	R	*AB*	2.62	1.9	74.3	17.0	69	3.1	31	18.1	0.17	77.6
2005	261	MV144	108	10/10/2005	1	E	*BB*	1.60	1.3	80.0	18.7	66	3.7	31	19.6	1.00	106.8
2005	262	Stoneyhurst	10	10/10/2005	2	R	*AB*	2.08	1.6	78.1	18.1	68	3.0	26	16.3	0.08	103.9
2005	263	Stoneyhurst	10	10/10/2005	2	E	*AB*	1.58	1.1	70.2	21.6	62	5.1	33	23.7	6.75	93.1
2005	264	Stoneyhurst	386/02	10/10/2005	2	.	*AA*	2.32	1.6	69.8	20.7	83	4.1	26	19.7	1.25	88.3
2005	265	Stoneyhurst	386/02	10/10/2005	2	.	*AB*	1.66	1.1	71.4	16.1	67	3.4	11	21.1	0.33	84.5
2005	267	Stoneyhurst	11	10/10/2005	1	E	*AC*	2.26	1.5	67.2	20.8	83	4.3	30	20.7	1.83	96.6
2005	273	Stoneyhurst	14	11/10/2005	2	E	*AC*	2.32	1.7	71.8	21.1	91	4.6	41	21.7	4.00	90.7
2005	274	MV144	117	11/10/2005	2	R	*AB*	2.48	1.8	74.1	17.2	84	4.4	32	25.4	0.83	84.2
2005	275	MV144	117	11/10/2005	2	E	*AB*	1.92	1.4	73.4	17.6	78	3.4	20	19.2	0.33	97.9
2005	279	Stoneyhurst	360/02	12/10/2005	2	E	*AA*	2.44	1.6	68.3	20.9	90	4.9	38	23.6	3.50	101.2
2005	281	Stoneyhurst	388/02	12/10/2005	2	E	*AB*	1.64	1.0	64.7	19.0	68	4.1	29	21.5	0.83	108.5
2005	282	Stoneyhurst	291/02	12/10/2005	2	R	*BC*	1.92	1.3	69.2	16.2	70	3.6	15	22.2	0.67	97.0
2005	283	Stoneyhurst	291/02	12/10/2005	2	R	*AB*	2.28	1.6	70.6	17.4	64	3.3	50	18.8	0.58	96.9
2005	284	Stoneyhurst	356/02	12/10/2005	2	E	*AB*	2.52	1.7	69.3	19.6	81	4.2	30	21.2	1.75	93.7
2005	285	Stoneyhurst	356/02	12/10/2005	2	E	*BC*	2.36	1.6	68.7	20.8	88	4.6	26	21.9	3.75	81.6
2005	286	Stoneyhurst	299/02	12/10/2005	1	R	*AC*	3.14	2.6	83.7	20.1	82	5.3	45	26.2	5.33	87.4
2005	288	MV144	107	12/10/2005	1	E	*AB*	2.68	1.7	66.7	16.0	73	3.1	23	19.5	0.17	92.4
2005	289	MV144	113	12/10/2005	1	E	*AB*	2.40	2.0	83.0	17.0	78	3.3	33	19.5	0.58	78.2
2005	290	MV144	124	12/10/2005	2	E	*AB*	2.14	1.6	76.1	18.6	83	3.7	31	20.1	0.58	98.3
2005	291	MV144	124	12/10/2005	2	R	*AB*	2.22	1.7	76.9	16.2	69	3.5	35	21.7	0.67	84.7
2005	292	Stoneyhurst	390/02	10/10/2005	1	R	*BB*	2.48	1.8	75.0	15.9	68	3.0	37	19.2	0.42	97.9
2005	294	Stoneyhurst	17	13/10/2005	2	R	*AC*	2.38	1.4	62.1	18.5	72	3.5	23	19.1	1.08	103.9
2005	296	MV144	131	14/10/2005	1	E	*AB*	1.88	1.4	75.4	16.0	67	3.4	26	21.3	0.33	92.1
2005	298	MV144	87	14/10/2005	2	E	*BB*	2.00	1.5	74.1	16.9	74	3.4	14	20.5	0.58	93.8
2005	299	Stoneyhurst	30	14/10/2005	2	.	*BC*	1.98	1.2	63.0	17.9	76	3.8	11	21.4	0.67	100.7
2005	300	Stoneyhurst	30	14/10/2005	2	.	*AA*	1.82	1.2	68.8	19.3	75	3.9	22	20.2	1.33	101.9
2005	301	Stoneyhurst	28	14/10/2005	1	R	*AA*	2.80	2.0	70.0	19.7	80	3.6	24	18.5	0.92	91.8
2005	302	Stoneyhurst	18	14/10/2005	1	R	*AA*	3.24	2.2	68.0	21.1	77	4.3	30	20.3	4.25	80.9
2005	304	MV144	65	14/10/2005	1	R	*AB*	2.28	1.8	81.3	16.3	89	3.0	12	18.5	0.17	91.0
2005	307	MV144	52	15/10/2005	1	E	*AB*	2.22	1.7	75.9	16.2	68	3.8	27	23.2	0.67	80.5
2005	309	Stoneyhurst	11	15/10/2005	1	R	*AB*	2.80	2.1	76.1	18.8	76	3.8	40	20.2	0.92	94.3
2005	310	MV144	132	15/10/2005	1	R	*BB*	2.46	1.9	78.7	18.3	82	3.1	49	17.0	0.58	96.6
2005	311	MV144	135	15/10/2005	1	E	*AB*	2.82	2.2	80.2	16.5	83	5.6	27	33.9	1.33	64.7
2005	313	MV144	136	15/10/2005	1	E	*AB*	2.28	1.7	75.2	15.8	64	3.9	31	24.5	0.42	85.5

2005	314	MV144	136	15/10/2005	2	E	*AB*	1.86	1.4	78.3	19.0	70	3.7	33	19.4	0.42	101.8
2005	316	Stoneyhurst	26	15/10/2005	1	E	*AC*	1.96	1.3	67.3	18.7	79	4.3	21	23.1	1.58	102.1
2005	322	Stoneyhurst	148	17/10/2005	1	E	*AA*	2.86	1.9	67.1	18.5	86	4.2	32	22.5	1.58	100.9
2005	323	Stoneyhurst	391/02	18/10/2005	1	R	*AB*	2.44	1.8	73.4	18.2	69	3.0	42	16.4	0.08	94.4
2005	324	Stoneyhurst	368/02	18/10/2005	1	E	*AB*	2.12	1.6	76.4	17.6	93	3.3	36	18.7	0.25	82.5
2005	325	MV144	126	18/10/2005	1	E	*AB*	2.72	2.2	83.0	16.8	74	3.6	46	21.4	0.50	83.7
2005	327	MV144	70	18/10/2005	1	E	*AB*	2.66	2.0	78.6	17.6	70	3.2	50	18.2	0.42	82.1
2005	329	MV144	75	18/10/2005	2	E	*BB*	2.00	1.5	76.9	18.8	92	3.5	37	18.5	0.58	85.8
2005	331	Stoneyhurst	38	18/10/2005	1	.	*AA*	2.38	1.7	71.9	19.4	88	3.7	16	19.2	0.83	83.5
2005	332	Stoneyhurst	19	18/10/2005	1	E	*BB*	2.88	2.3	82.4	16.3	82	2.9	33	17.6	0.17	74.8
2005	334	Stoneyhurst	5	19/10/2005	2	E	*AB*	1.88	1.2	66.7	18.0	82	3.7	21	20.3	1.08	95.8
2005	335	Stoneyhurst	5	19/10/2005	2	R	*BC*	1.94	1.4	71.2	19.1	81	4.2	33	21.8	1.50	92.1
2005	337	Stoneyhurst	379/02	19/10/2005	1	R	*AB*	2.76	2.0	75.3	19.2	79	4.9	36	25.3	1.58	85.4
2005	338	Stoneyhurst	32	19/10/2005	1	R	*AC*	3.00	2.0	67.0	20.7	65	3.2	45	15.5	1.00	99.5
2005	340	MV144	61	20/10/2005	1	E	*AB*	2.64	2.0	75.2	14.6	82	3.5	24	24.3	0.83	78.7
2005	341	Stoneyhurst	393/02	20/10/2005	1	R	*AC*	2.86	2.0	70.2	19.9	82	3.6	10	18.3	1.67	91.8
2005	342	MV144	143	20/10/2005	1	R	*AB*	2.26	1.8	81.2	16.3	87	2.9	23	17.9	0.33	84.5
2005	343	Stoneyhurst	298/02	20/10/2005	1	R	*AB*	2.26	1.6	70.6	18.4	83	3.7	11	19.9	0.50	95.1
2005	345	Stoneyhurst	365/02	20/10/2005	2	R	*AB*	2.44	2.0	75.1	19.3	71	5.3	21	27.3	4.08	79.6
2005	348	Stoneyhurst	355/02	20/10/2005	1	R	*AB*	2.58	1.8	73.5	18.2	77	3.3	36	18.2	0.33	99.1
2005	349	MV144	76	21/10/2005	1	E	*AB*	2.68	2.0	77.9	18.7	68	4.1	30	21.9	1.33	89.7
2005	350	Stoneyhurst	31	20/10/2005	1	R	*AC*	2.78	1.8	65.3	20.2	70	4.4	35	21.8	3.08	83.8
2005	351	MV144	142	22/10/2005	1	R	*BB*	2.26	1.8	79.6	15.5	73	3.2	45	20.5	0.42	82.9
2005	352	Stoneyhurst	383/02	22/10/2005	1	.	*AB*	2.74	2.0	75.3	18.0	83	3.7	17	20.6	0.83	85.8
2005	354	MV144	118	22/10/2005	2	E	*AB*	2.66	2.0	75.5	16.0	87	3.1	20	19.2	0.33	94.5
2005	355	Stoneyhurst	21	23/10/2005	1	R	*AC*	2.24	1.8	80.2	17.5	77	3.1	33	17.5	0.17	84.2
2005	356	Stoneyhurst	361/02	25/10/2005	2	R	*BC*	2.30	1.6	71.6	18.2	74	4.2	36	23.2	1.67	95.7
2005	357	Stoneyhurst	361/02	25/10/2005	2	E	*AB*	2.20	1.4	64.7	17.7	77	4.7	18	26.4	2.33	97.3
2005	358	Stoneyhurst	4	25/10/2005	1	E	*AC*	2.28	1.5	70.2	20.4	70	4.1	39	20.2	2.17	101.4
2005	359	MV144	93	25/10/2005	1	.	*AB*	2.12	1.6	78.3	17.5	68	2.8	31	15.8	0.33	75.7
2005	360	MV144	130	26/10/2005	1	E	*AB*	2.34	1.8	79.2	18.0	86	2.8	26	15.8	0.33	78.0
2005	363	Stoneyhurst	383/02	28/10/2005	2	E	*BC*	2.08	1.6	77.8	17.0	84	3.2	42	18.6	0.50	89.0
2005	366	Stoneyhurst	8	4/11/2005	1	R	*AB*	2.76	2.3	83.4	18.3	85	3.6	52	19.7	1.17	75.4
2005	367	Stoneyhurst	144	2/11/2005	1	E	*AC*	2.40	1.6	65.9	19.4	82	4.7	16	24.2	2.75	89.8
2009	513	YT Southdown	822	.	2	R	*BB*	2.1	1.5	71.9	19.5	80	4.5	29	23.1	2.55	93.3
2009	514	sire 2	590	.	2	R	*AB*	2.6	2	76.6	25.1	90	6.2	15	24.7	20.5	73.8
2009	516		509	.	2	R	*CC*	2.4	1.5	62.2	23.5	88	5.7	16	24.1	11.5	84.8
2009	518	sire 2	555	.	1	R	*AA*	2	1.5	76.2	20.1	73	4.1	22	20.4	1.1	79.9
2009	521	sire 2	591	.	1	R	*AB*	2.6	1.9	74.3	21.7	76	5.6	16	26	8.8	91.8
2009	523	sire 2	592	.	2	R	*AB*	2.2	1.6	70.8	19.6	93	4.6	20	23.5	3.25	101.6

2009	524	sire 2	616	.	1	R	*AA*	1.8	1.2	66.6	17.7	78	4.6	13	26.2	2.05	101.2
2009	525	sire 2	564	.	2	R	*AB*	1.8	1.3	74.8	18.3	66	3.7	14	20.5	0.8	98.2
2009	528	YT Southdown	427	.	1	R	*BB*	2.4	1.7	69.6	21.8	82	4.8	14	22	5.4	96.4
2009	529	YT Southdown	355	.	1	R	*AA*	1.7	1	61.1	19.6	71	4.3	12	21.9	1.55	129.7
2009	530	YT Southdown	822	.	2	R	*BB*	2.1	1.4	66.8	18.5	91	4.2	16	22.5	1.2	106.6
2009	532	YT Southdown	1505	.	2	R	*AA*	2.1	1.4	67.6	17	93	3.3	28	19.2	0.35	93.4
2009	533	YT Southdown	1096	.	1	R	*AB*	2.1	1.4	66.1	20.3	62	4.3	39	21.1	1.4	109.4
2009	535	YT Southdown	2026	.	2	R	*BB*	2.2	1.6	70.9	17.5	100	3.7	18	21.4	0.45	78.3
2009	536	YT Southdown	2610	.	2	R	*BB*	2.4	1.6	68.6	20	83	4.9	24	24.7	2.95	88.5
2009	537	YT Southdown	218	.	1	R	*AA*	2.1	1.5	69.6	20.9	75	4.9	18	23.5	3.9	106.8
2009	539	YT Southdown	222	.	1	R	*AA*	2.5	1.9	74.4	20.8	80	4.5	22	21.9	3.35	82
2009	540	YT Southdown	1488	.	1	R	*AB*	2.1	1.4	67.9	20.1	74	3.8	24	18.9	0.85	112.9
2009	541	YT Southdown	63	.	2	R	*AA*	2.6	1.7	63.8	21.8	79	5.2	27	23.9	5.2	71.3
2009	544	YT Southdown	6	.	2	R	*AB*	2.9	2.3	78.7	18.3	94	4	28	21.7	0.75	79.9
2009	546	YT Southdown	554	.	1	R	*AA*	2	1.5	75.4	17.6	79	4.1	28	23.3	1.1	104.9
2009	547	YT Southdown	1137	.	2	R	*AA*	2	1.4	68.8	20.2	78	5.3	18	26.1	5.1	97.4
2009	548	YT Southdown	2206	.	2	R	*BB*	2.1	1.4	66.6	20.8	72	4.1	35	19.9	1.5	95.9
2009	549	YT Southdown	448	.	2	R	*AB*	2	1.3	64.7	18	65	3.4	25	18.9	0.25	106.3
2009	550	YT Southdown	2009	.	2	R	*BB*	1.8	1.2	69.4	18.5	68	4.3	21	23.3	1.6	125.6
2009	551	sire 2	1182/04	.	1	R	*AB*	2.6	1.8	67.9	22.4	89	4.9	16	22	6.8	79.8
2009	552	sire 2	519	.	2	R	*AA*	2.2	1.6	73.2	18.5	89	4.3	27	23.3	1.1	68.7
2009	553	sire 2	454	.	1	R	*AA*	2.6	1.8	71	18	79	3.8	23	21.1	0.7	84.4
2009	554	sire 2	1191/04	.	1	R	*AA*	1.9	1.2	65.8	22	71	4.5	12	20.3	4.15	99
2009	555	sire 2	573	.	2	R	*AB*	2.5	1.7	69.5	21.9	87	5.2	17	23.6	7.1	93.6
2009	556	sire 2	531	.	2	R	*AB*	2.6	1.6	63.1	19.9	96	4.8	12	24	2.4	111.1
2009	557	sire 2	564	.	2	R	*BB*	2.2	1.6	72.1	19.6	71	4.7	10	24.1	2.75	112.1
2009	558		531	.	2	R	*CC*	2.2	1.4	62.6	20.6	80	4.2	18	20.4	2.55	127.9
2009	562	sire 2	1192/04	.	2	R	*AB*	1.8	1.1	62.9	16.7	68	3.2	24	19	0.3	107.7
2009	563	sire 2	650	.	2	R	*BB*	1.8	1.3	72.4	17	85	4	24	23.6	0.1	86.3
2009	564	sire 2	261	.	1	R	*AB*	2.5	1.6	64.6	21	72	4.4	17	20.9	2.15	90.2
2009	565	sire 2	462	.	1	R	*BB*	1.5	1.1	72.6	20.9	68	4.7	23	22.5	3.6	95.2
2009	566	sire 2	266	.	1	R	*AA*	2.4	1.6	67.4	19.6	70	4.1	11	20.9	1.05	85.4
2009	567	sire 2	313	.	1	R	*BB*	1.8	1.2	68.2	18.5	76	4.1	14	21.9	1.25	107.8
2009	568	sire 2	268	.	2	R	*AA*	2.1	1.6	77.7	20.1	75	4.6	37	22.7	3	82.4
2009	569	sire 2	357/02	.	1	R	*AB*	2	1.4	69.7	18.9	57	4.4	34	23.5	1.65	90
2009	570	sire 2	489	.	1	R	*BB*	2.2	1.5	67	19.9	73	4.4	7	22.2	1.7	106.6
2009	572	sire 2	369/02	.	1	R	*AB*	2.5	1.9	75.2	18.9	85	4.4	27	23.2	1.15	82.8
2009	573	sire 2	480	.	1	R	*BB*	2.2	1.5	68.2	20.2	76	4.7	9	23.4	2.45	96.4
2009	574	sire 2	263	.	1	R	*AB*	2	1.5	72.7	19.5	77	4.6	14	23.8	2.2	95.2
2009	578	YT Southdown	586	.	2	R	*AB*	2.1	1.5	70.5	21.6	52	5.3	16	24.7	6.5	102.9

2009	579	YT Southdown	6/07	.	1	R	*AB*	1	0.7	71.7	20.9	49	4.8	21	22.8	3.4	139.7
2009	582	YT Southdown	1449	.	1	R	*BB*	2.4	2	82.9	18.5	88	3.8	26	20.5	0.6	112.1
2009	583	YT Southdown	69	.	1	R	*BB*	2.6	1.5	59.2	20.9	61	5	14	23.9	5.35	106.8
2009	585	YT Southdown	850	.	2	R	*AB*	2.1	1.3	60.1	20.1	65	4.5	12	22.2	1.75	114.5
2009	586	YT Southdown	435	.	2	R	*BB*	2.4	1.6	68.2	20	73	4.1	21	20.5	0.6	118.8
2009	587	YT Southdown	2591	.	2	R	*BB*	2.1	1.4	69	20.3	78	4.8	15	23.5	2.55	117.3
2009	588	YT Southdown	2591	.	2	R	*BB*	2.3	1.4	59.7	22.2	69	5	24	22.5	5.45	146.7
2009	589	YT Southdown	14/07	.	1	R	*AA*	1.6	1.1	69.2	19.9	72	4.5	15	22.8	1.4	82.8
2009	591	YT Southdown	1341	.	1	R	*AB*	1.4	1.1	76.6	17.5	74	3.3	26	19.1	0.15	107.3
2009	592	YT Southdown	1609	.	2	R	*BB*	2.4	1.4	57.7	22.3	63	4.5	29	20.1	4.15	78.6
2009	593	YT Southdown	665	.	1	R	*BB*	1.8	1.5	80.8	18.7	75	4.7	42	25.2	2.1	78
2009	594	YT Southdown	47/07	.	1	R	*AB*	2	1.4	70.9	23.5	78	7	25	29.6	17.6	102.5
2009	597	YT Southdown	1281	.	1	R	*AB*	2.3	1.5	66.7	21.5	74	4.7	23	21.6	3.3	129.6
2009	598	YT Southdown	2095	.	1	R	*AB*	2	1.3	63.9	19.9	83	4.3	15	21.8	1.25	109.3
2009	602	sire 2	1186/04	.	1	R	*AA*	2.1	1.4	65.8	21.2	60	4.9	17	23.2	3.85	108.3
2009	603	sire 2	514	.	2	R	*AA*	2	1.4	71.6	18	78	4.5	21	25.3	1	80
2009	604	sire 2	280	.	1	R	*BB*	1.9	1.3	68.2	20.2	59	4.6	22	23	2.9	97.1
2009	606		594	.	1	R	*CC*	2.2	1.6	74.4	21.8	69	4.8	17	21.9	4.1	121.2
2009	607	sire 2	331	.	1	R	*AA*	2.7	1.9	71.8	16.8	79	3.8	22	22.4	0.6	81.3
2009	608	sire 2	335	.	1	R	*AA*	2.5	1.9	74.1	18.1	84	4.4	17	24.5	1	100.7
2009	609	sire 2	279	.	1	R	*AA*	1.8	1.3	72.7	18.8	68	4.5	24	24.1	1.25	80.6
2009	610	sire 2	282	.	2	R	*BB*	1.9	1.4	75.8	18.1	82	4.2	19	23	0.55	73.6
2009	611	sire 2	282	.	2	R	*AB*	1.8	1.2	64.3	17.7	78	3.8	14	21.6	0.6	111.1
2009	612	sire 2	492	.	1	R	*BB*	1.9	1.5	77	17.1	73	4.1	21	23.8	0.7	100
2009	613	sire 2	483	.	1	R	*BB*	2.1	1.3	63.5	20.2	68	4.5	21	22	2.3	93.2
2009	614	sire 2	402	.	1	R	*AB*	1.8	1.4	75.1	18.4	70	4.4	28	24	1.65	84.4
2009	770	YT Southdown	256	.	1	E	*BB*	2.6	2	77.3	18.5	112	3.4	25	18.4	0.25	74.9
2009	776	YT Southdown	2622	.	1	E	*AB*	2.2	1.5	67.6	21	72	4.5	25	21.5	2.95	89.8
2009	777	YT Southdown	1903	.	1	E	*BB*	2.3	1.7	74.4	20.8	88	4.8	28	22.8	3.95	95.3
2009	778	YT Southdown	2168	.	1	E	*BB*	2	1.6	80	17.5	73	4.6	22	26.1	1.2	95.3
2009	780	YT Southdown	1684	.	1	E	*BB*	2.5	1.7	66.4	20.7	76	5.5	15	26.4	4.9	89.5
2009	781	YT Southdown	2206	.	2	E	*BB*	2.1	1.4	65.8	22.1	70	5.6	30	25.2	8	92.9
2009	783	YT Southdown	713	.	1	E	*BB*	2	1.5	73.1	19.4	75	4.5	17	23.4	0.95	88.8
2009	784	YT Southdown	1137	.	2	E	*BB*	2	1.3	64.9	17.7	93	4	24	22.5	0.4	100.9
2009	785	YT Southdown	2343	.	1	E	*BB*	2.6	1.8	69	19.6	93	4.3	25	21.9	1.5	109.1
2009	786	YT Southdown	1505	.	2	E	*BB*	2.1	1.4	64.3	20.4	72	4.7	38	23.2	3.25	99
2009	787	YT Southdown	157	.	1	E	*BB*	2	1.2	61.9	22.2	77	4.6	15	20.9	4.2	105.1
2009	788	YT Southdown	2026	.	2	E	*BB*	2.4	1.4	58	20.4	72	5.1	16	25	4.85	119
2009	789	YT Southdown	2610	.	2	E	*BB*	1.8	1.2	68.5	19.2	61	4.3	14	22.4	1.85	117.2
2009	790	YT Southdown	611	.	1	E	*AB*	1.5	1.2	82.4	17.9	74	4	20	22.6	0.35	82.3

2009	791	YT Southdown	121	.	1	E	*AB*	2.1	1.2	56.3	20.5	64	4.8	16	23.3	3.1	111.2
2009	801	YT Southdown	740	.	1	E	*AA*	2.4	1.7	71.4	20.8	96	5.1	20	24.6	5.75	93.8
2009	802	YT Southdown	2009	.	2	E	*AA*	1.8	1.2	64.4	20.6	76	4.1	26	19.7	1.85	104.9
2009	804	YT Southdown	448	.	2	E	*AB*	1.7	1	59	18.5	84	3.6	22	19.7	0.6	113
2009	805	YT Southdown	1266	.	1	E	*AB*	2.3	1.7	72.8	19.9	88	4.3	24	21.4	1.95	79.7
2009	807	YT Southdown	586	.	2	E	*AB*	2.6	1.8	70.4	19.7	97	4.3	17	21.8	1.4	83.5
2009	814	YT Southdown	522	.	2	E	*AB*	1.9	1.2	63.9	19.3	71	4.2	39	21.7	1.7	102
2009	815	YT Southdown	453	.	1	E	*AA*	2.3	1.6	67.7	20.6	89	4.4	27	21.5	2.3	83.5
2009	819	YT Southdown	850	.	2	E	*AB*	2.4	1.6	66.8	20.3	90	5	17	24.8	2.4	99.9
2009	821	YT Southdown	2498	.	1	E	*BB*	2.2	1.5	70.3	20.8	63	4.1	31	19.6	1.85	108
2009	822	YT Southdown	435	.	2	E	*AA*	2.4	1.9	79.1	20.4	84	4.7	35	22.9	1.6	79.4
2009	823	YT Southdown	373 Pink	.	1	E	*BB*	2.6	1.6	63	21.1	93	4.5	19	21.2	3.25	115.6
2009	826	YT Southdown	2147	.	2	E	*AA*	2.1	1.4	67.1	19	74	4.8	12	25.5	2.45	102.3
2009	827	YT Southdown	1079	.	1	E	*AB*	2.6	1.8	70.7	21.4	78	5	24	23.2	4.7	103.9
2009	828	YT Southdown	1609	.	2	E	*BB*	1.7	1.1	64.1	18.9	90	4.5	15	24	1	110
2009	830	YT Southdown	28/07	.	1	E	*AB*	1	0.7	70.6	21.1	49	5.1	32	24.4	4.05	94.1
2009	832	YT Southdown	1901	.	1	E	*AB*	2.6	1.9	73.1	20.3	89	4.5	30	22.3	1.4	71.7
2009	833	YT Southdown	182	.	2	E	*BB*	2.7	2.1	77	19.9	104	4.5	22	22.8	2.1	87.1
2009	834	YT Southdown	182	.	2	E	*BB*	2.4	1.6	65.8	18.3	100	3.8	21	20.8	1.05	95.9
2009	835	YT Southdown	525	.	2	E	*AB*	2.5	1.8	73.6	19.9	96	3.9	24	19.6	0.9	98.4
2009	836	YT Southdown	525	.	2	E	*BB*	2.6	1.9	72.8	20.8	91	4.7	16	22.6	3.15	92
2009	837	YT Southdown	316	.	1	E	*AB*	2	1.2	61.7	19.5	76	4.1	30	21.2	1.05	144.2
2009	838	YT Southdown	255	.	2	E	*BB*	1.8	1.2	64.1	19.6	77	3.9	18	20.1	0.4	127.9
2009	839	YT Southdown	255	.	2	E	*BB*	2.1	1.5	71.3	19.8	86	4.7	24	24	2.8	88.5
2009	841	YT Southdown	109	.	2	E	*BB*	2.1	1.4	68.1	18.7	86	4.6	8	24.9	1.5	89.2
2009	842	YT Southdown	109	.	2	E	*BB*	2.3	1.5	66.6	20.3	76	4.6	15	22.9	2.35	82.9
2009	843	sire 2	1055	.	2	E	*BB*	2.8	2.1	73.9	18.7	80	4.4	18	23.5	1.2	86
2009	844	sire 2	1055	.	2	E	*AA*	2.6	1.8	67.3	22.7	82	6.1	18	26.8	11.25	92.5
2009	845	sire 2	575	.	1	E	*BB*	2.4	1.7	68.8	19.8	103	4	19	20.4	1.05	87.9
2009	847	sire 2	569	.	1	E	*BC*	2.3	1.6	68.3	22.9	63	5.4	22	23.6	7.55	90
2009	850	sire 2	1029	.	1	E	*AB*	2.5	1.8	70.6	21.8	86	5.3	34	24.1	7.25	90.1
2009	852	sire 2	597	.	1	E	*AA*	1.4	1	71.9	20.8	88	5.4	12	25.7	4.8	91.8
2009	853	sire 2	592	.	2	E	*BC*	2.2	1.4	65	23.6	91	6.3	18	26.7	14.6	95.1
2009	854	sire 2	571	.	2	E	*AA*	2.9	2.1	73.3	18.9	78	3.6	23	19	0.35	112.3
2009	855	sire 2	571	.	2	E	*AB*	2.1	1.2	56.7	21.8	60	5.2	14	23.8	7.4	100
2009	857	sire 2	548	.	1	E	*AB*	1.5	1.1	72.4	21.5	61	5.5	20	25.5	6.6	112
2009	858	sire 2	460	.	1	E	*AB*	1.8	1.3	70	19	83	3.9	28	20.6	0.7	97.7
2009	859	sire 2	650	.	2	E	*BB*	1.6	1.3	83	18	72	5.1	24	28.2	1.1	84.2
2009	860	sire 2	384/02	.	1	E	*BB*	2.2	1.4	65.5	21.9	78	5	16	22.9	6.8	91.9
2009	861	sire 2	464	.	1	E	*AB*	2.2	1.5	68.5	21.3	75	5.3	17	25	5.75	106.6

2009	862		596	.	1	E	*CC*	2.5	1.8	72.4	20.7	80	4.1	18	20	2.3	76.9
2009	863	sire 2	405	.	1	E	*AA*	2.6	1.8	68.2	19.1	77	4.4	19	23.1	1.5	102.2
2009	865	sire 2	573	.	2	E	*AA*	2.1	1.5	70.9	20.3	85	4.7	21	23	2.8	91.2
2009	866	sire 2	268	.	2	E	*AB*	2.3	1.8	77.3	19.6	82	4.9	25	25	3.25	92.8
2009	867	sire 2	563	.	1	E	*BC*	2	1.4	69.5	21.8	74	5.9	18	26.9	9.3	99.1
2009	868	sire 2	286	.	1	E	*BB*	2	1.4	70.1	19.7	70	4.7	21	24.1	2.2	84.9
2009	869	sire 2	272	.	1	E	*AA*	2.2	1.6	72.7	20.5	76	4.7	28	22.7	3.15	88
2009	871	sire 2	601	.	1	E	*AB*	2.4	1.8	77.3	19.3	85	4.2	29	21.5	0.85	77.8
2009	874	sire 2	2141	.	1	E	*AB*	2.3	1.4	59.8	20.9	75	5.2	22	24.7	4.15	127.7
2009	876	sire 2	404	.	1	E	*AA*	2.3	1.7	74.2	18.2	88	5	24	27.6	2.6	76.8
2009	877	sire 2	514	.	2	E	*AA*	2.2	1.7	75.9	19.3	89	4.3	21	22.2	1.55	76.8
2009	878	sire 2	515	.	1	E	*BB*	2.5	2	81.9	21.9	86	5.3	40	23.9	5.4	80.1
2009	879	sire 2	275	.	2	E	*BB*	1.8	1.1	63.3	17.1	83	3.8	20	22.4	0.75	130.1
2009	880	sire 2	458	.	1	E	*AA*	2	1.3	67.1	22.5	63	4.8	22	21.3	6	91.6
2009	882	sire 2	508	.	2	E	*AB*	.		76.5	21	103	5	19	24	4.85	69.1
2009	884	sire 2	468	.	1	E	*BB*	1.9	1.3	69.5	18.5	74	4.4	24	23.7	1.8	98.6
2009	885	sire 2	265	.	1	E	*AA*	2.2	1.6	71.3	20.2	78	4.6	20	22.6	2.15	96.4
2009	886	sire 2	576	.	1	E	*BB*	2	1.4	70.7	16.8	84	4.8	18	28.6	1	76.5
2011	4003	Southdown	44/07		1	R	*AB*	2.12	1.8	79.8	15.3	94	3.2	7	21	0.15	63.3
2011	4004	Merino	563		1	R	*BC*	2.38	1.7	69.6	22.6	71	5.5	19	24.4	8	94.2
2011	4011	Merino	.		1	E	*AA*	2.32	1.7	71.6	16.3	85	3.3	14	20.5	0.2	63.2
2011	4012	Southdown	26/07		1	R	*AB*	2.06	1.7	81.8	17.3	89	3.7	28	21.2	0.55	64.2
2011	4013	Merino	769		1	E	*AA*	2.48	1.7	69.2	19.2	100	3.3	21	16.9	0.3	80.6
2011	4016	Southdown	548		1	R	*AA*	2.66	1.9	69.9	21.6	82	4.8	23	22.4	5.55	86.4
2011	4020	Merino	886		1	E	*AA*	1.88	1.4	72.1	16.7	69	3.7	17	22.3	0.5	80.1
2011	4024	Merino	787/09		1	R	*AB*	1.38	1	69.6	15.7	67	3.8	7	24	0.15	59.5
2011	4032	Merino	854		1	R	*BB*	2.26	1.7	73.4	16.7	84	3.3	19	19.8	0.5	92.5
2011	4033	Merino	788		1	E	*AB*	2.58	1.8	71.5	16.3	66	2.9	20	18.1	0.2	65.5
2011	4034	Merino	877		1	R	*BB*	2.7	1.9	70.2	17.7	82	3.5	22	19.8	0.4	135.9
2011	4036	Merino	865/09		1	E	*AA*	2.46	1.8	72.5	18.9	92	3.7	25	19.7	0.75	85.3
2011	4038	Merino	732/08		1	R	*BB*	2.92	2.4	78.6	16.7	94	3.6	17	21.4	0.4	62.8
2011	4039	Southdown	161		1	R	*AA*	1.94	1.3	70.1	18.8	75	3.6	23	19.1	0.75	103.5
2011	4040	Merino	814		2	E	*AA*	2.94	2.6	87.3	18	102	2.9	23	16	0.2	56.5
2011	4043	Merino	815		1	R	*AB*	2.4	1.9	79.3	16.5	67	2.6	25	16	0.2	67.4
2011	4044	Merino	781		1	R	*BB*	2.32	1.8	77.5	17.8	91	3.6	26	20.2	0.5	75
2011	4045	Merino	6/07		3	E	*AB*		1.1	78.2	20.2	92	4	28	19.7	1.15	127.9
2011	4046	Merino	6/07		3	E	*AB*	1.92	1.6	82.6	18	89	3.5	17	19.4	0.25	80.1
2011	4047	Merino	6/07		3	E	*AB*	2.34	2	85.1	17.2	97	3.4	28	20	0.5	81.7
2011	4048	Merino	917/08		1	E	*AB*	2.83	2.4	84.8	18.1	114	3.7	12	20.3	0.5	83.1
2011	4050	Merino	836		1	R	*BB*	3.3	2.6	79.3	16.9	110	3	19	18	0.4	76.6

2011	4051	Merino	868/09	2	E	*BB*			71.9	16.4	88	3.1	40	18.9	0.2	88
2011	4052	Merino	868/09	2	R	*AA*	2.24	1.7	77.2	17.8	80	3.3	22	18.6	0.25	63.2
2011	4053	Merino	816	2	E	*AA*	1.74	1.4	76.6	17.2	79	4.2	12	24.4	1.1	75
2011	4057	Merino	NT	2	E	*AA*	2.88	1.8	63.3	16.7	96	4	14	23.9	0.7	76.4
2011	4058	Merino	NT	2	E	*AA*	3.1	2	65.4	16.6	104	3.9	11	23.6	0.65	55.1
2011	4059	Merino	50/07	1	R	*AA*	3.36	2.6	76.8	16.6	98	3.2	22	19.2	0.25	60.7
2011	4065	Merino	784/09	1	R	*AB*	3.06	2.4	81.3	18.4	102	3.4	29	18.6	0.3	55.4
2011	4066	Merino	763	1	E	*AB*	2.48	1.9	76.8	18.6	95	3.4	22	18.6	0.5	67.2
2011	4067	Merino	806/09	1	E	*AB*	2.96	2.3	75.8	18.1	104	2.9	32	15.9	0.4	65.7
2011	4068	Merino	789	1	E	*BB*	2.44	1.7	70.5	17.1	91	2.8	24	16.4	0.1	99
2011	4069	Merino	855/09	1	E	*AB*	2.48	1.6	64.7	16.4	75	4.1	13	24.9	0.7	99.8
2011	4070	Merino	802	1	E	*BB*	2.58	1.9	74.5	16.7	90	3.6	18	21.7	0.45	105.3
2011	4071	Merino	779	1	R	*AA*	2.8	2.1	89.8	16	90	3.1	30	19.2	0.25	42.3
2011	4072	Merino	780/09	1	E	*AA*	1.98	1.6	82.3	17.7	97	3.4	19	19.1	0.3	74.7
2011	4074	Merino	879/09	1	E	*BB*	2.54	1.7	68.5	18.2	97	3	21	16.3	0.45	97.8
2011	4075	Merino	827/09	1	R	*AB*	2.56	2	78.4	18.1	94	3	37	16.4	0.25	82.6
2011	4076	Merino	778/09	1	R	*BB*	2.68	2.1	78.1	17	94	3.1	39	18.1	0.3	73.4
2011	4077	Merino	882	1	R	*AB*	2.6	2.2	85.1	16.8	103	2.7	28	16.4	0.3	49
2011	4078	Merino	832/09	2	R	*AB*	2.94	2.2	72.2	17.2	99	3	38	17.4	0.2	59
2011	4079	Merino	863	1	E	*BB*	2.78	2	72.3	18.1	94	3.6	22	20	0.6	78.9
2011	4081	Merino	781	1	R	*AB*	2.6	2.2	86.2	17.1	96	3.4	22	20.1	0.5	70.8
2011	4082	Merino	819/09	2	E	*BB*	2.62	1.9	74.1	17.4	98	3.5	27	20	0.6	73.6
2011	4083	Merino	819/09	2	R	*BB*	2.14	1.8	86.8	16.9	79	4.1	15	24.4	0.8	88.5
2011	4084	Merino	821/09	1	E	*AA*	2.88	1.7	72.6	18.3	92	3.3	20	18.2	0.2	86.4
2011	4085	Merino	770/09	2	E	*AB*	2.42	1.7	70.7	17	104	3	24	17.7	0.45	78.4
2011	4087	Merino	808/09	1	R	*BB*	2.18	1.6	73.6	18	95	3.6	12	19.8	0.3	87.9
2011	4088	Merino	783/09	1	R	*BB*	2.72	2.3	84.6	16.8	84	3.5	22	20.6	0.5	90.1
2011	4089	Merino	757/08	1	R	*AB*	2.04	1.4	72.1	16.3	76	3.2	11	19.4	0.2	85.7
2011	4091	Merino	885/09	1	R	*AA*	2.5	1.8	73.7	17.3	81	3.4	38	19.8	0.4	85.5
2011	4094	Merino	804/09	1	R	*BB*	2.76	2.1	75	16.7	73	3.5	15	20.6	0.45	82.5
2011	4096	Merino	874/09	1	E	*AB*	3.02	2.6	85.5	17.3	94	3.4	41	19.7	0.5	69.8
2011	4097	Merino	850/09	1	E	*AA*	2.84	2.1	72.1	17.3	89	3.8	31	22	0.9	65
2011	4098	Merino	725/08	1	R	*AB*	2.76	2.1	75.5	16.2	86	3.3	14	20.4	0.1	45.5
2011	4099	Merino	815	2	E	*AA*	2	1.5	74.7	18.8	95	3.2	26	17.1	0.15	84.1
2011	4100	Merino	815	2	E	*AA*	2.62	1.9	70.9	17.3	93	3.1	15	17.7	0.15	71.8
2011	4104	Merino	875	1	E	*BB*	2.4	1.9	77.3	19.4	87	3	34	15.4	0.35	85.5
2011	4105	Merino	95/07	1	R	*AA*	2.3	1.6	71.4	14.6	80	2.9	18	19.9		62.1
2011	4108	Merino	823	1	R	*BB*	2.88	2.2	76.7	19.8	102	3.7	40	18.6	1.05	84
2011	4109	Merino	794/08	1	E	*AA*	2.78	2	72.3	15.9	103	3.4	16	21.1	0.2	55.3
2011	4110	Merino	818	1	E	*AB*	2.74	2	72.8	16.7	73	3	27	17.8	0.15	77.1

2011	4111	Merino	820/09	1	E	*AA*	2.58	2.3	88.9	19.8	89	3.6	36	18.4	0.55	76.3
2011	4113	Merino	857	1	E	*AB*	2.28	1.9	83.9	16.3	98	3.8	9	23	0.1	84.2
2011	4114	Merino	854	1	E	*AB*	2.48	1.9	76.4	18.5	99	3.1	19	17	0.25	74.4
2011	4115	Merino	.	2	R	*AB*	2.36	2	82.1	16.4	90	2.9	21	18	0.55	66.8
2011	4116	Merino	.	2	E	*AB*	1.98	1.6	79.2	18	101	3.2	38	17.6	0.4	59.6
2011	4117	Merino	10/07	1	E	*AA*	2.84	2.3	80.3	18.6	78	4	13	21.4	0.7	67.1
2011	4118	Merino	791	2	E	*BB*	3.02	2.5	82.5	18.2	96	3.4	26	18.5	0.65	73.8
2011	4119	Merino	791	2	E	*AA*	2.08	1.6	78.3	15.2	95	3.8	15	24.8	0.45	65.2
2011	4120	Merino	545	1	E	*BB*	2.04	1.2	59	24.9	61	5.6	9	22.4	17.15	111.4
2011	4121	Merino	822/09	1	R	*BB*	3.16	2.3	72.6	16.3	95	3.2	11	19.6	0.2	63.6
2011	4122	Merino	869/09	1	E	*AB*	1.48	1.3	86.4	16.5	95	3.6	8	22.1	0.05	57.4
2011	4124	Merino	779	1	E	*AA*	3.54	2.6	72.6	18	114	3.4	19	19	0.55	74.3
2011	4125	Merino	14/07	2	R	*AA*	2.32	1.6	71.1	17.9	84	3.4	33	19	0.45	72.8
2011	4126	Merino	14/07	2	E	*AB*	2.46	1.8	71.6	17	86	3	33	17.6	0.4	83.4
2011	4127	Merino	42/07	1	E	*BB*	2.88	2.2	75.6	18.1	104	3.5	28	19.4	0.6	87.7
2011	4129	Southdown	570	2	E	*AB*	2.1	1.6	75.9	20.4	83	4.1	37	20.1	1.25	80.6
2011	4132	Merino	830	1	E	*AB*	2.96	2.1	70.6	19.9	88	3.9	25	19.3	1.55	90.1
2011	4133	Merino	776	2	R	*AB*			69.6	19.7	81	3.6	31	18.5	0.85	119.1
2011	4134	Merino	776	2	R	*AB*	2.12	1.6	74.4	15.3	87	2.8	38	18	0.2	55.4
2011	4136	Merino	842/09	1	E	*AB*	2.64	2.4	88.2	17.9	103	3.4	14	18.8	0.7	57.9
2011	4138	Merino	791	1	R	*AA*	2.48	2	78.6	17.9	90	3.8	26	21.3	0.8	71.7
2011	4140	Merino	758	2	R	*AB*	2.92	2.1	71.2	19.4	103	3.6	31	18.6	0.7	94.7
2011	4141	Southdown	218/08	1	E	*AB*	2.26	1.9	69.1	18.3	74	3.1	30	17	0.2	109.2
2011	4142	Southdown	399	1	R	*AB*	2.06	1.3	65.1	17.5	90	3.4	31	19.5	0.6	93.2
2011	4143	Southdown	3159	1	R	*AB*	2.64	1.9	70.7	21.4	81	4.2	36	19.4	2.25	113.5
2011	4145	Merino	786/09	1	E	*AA*	2.38	1.8	75.2	17.6	76	4	24	22.6	0.8	84.9
2011	4146	Merino	807/09	1	R	*AB*	2.5	2	78.6	17.3	92	3.4	10	19.9	0.45	100.7
2011	4147	Merino	854	1	R	*AB*	3.22	2.5	74.9	17.6	91	3.5	21	19.7	0.55	74.6
2011	4149	Merino	816	1	R	*BB*	2.8	2.1	73.3	17.7	96	3.3	24	18.8	0.05	64.3
2011	4151	Merino	845/09	2	E	*AA*	2.64	1.8	65	17.8	100	3.3	26	18.3	0.45	63.5
2011	4152	Merino	835	2	E	*AA*	2.4	1.7	69.5	18.4	89	3.7	34	20.2	1.45	106.1
2011	4153	Merino	835	2	E	*BB*	2.8	2.1	75.1	17.9	102	4.6	19	25.5	1.9	64.9
2011	4154	Merino	878	1	R	*AB*	2.24	1.7	74.6	17.7	123	2.9	16	16.2	0.2	91.8
2011	4155	Merino	835	1	E	*BB*	3.18	1.9	59.5	19.5	115	3.5	41	18.1	0.75	108.4
2011	4157	Merino	802	1	E	*AA*	2.5	1.9	77.8	16.7	86	2.9	21	17.4	0.35	63.5
2011	4158	Merino	771	1	E	*AB*	2.7	2.2	78.5	17.3	116	3.2	22	18.3	0.2	48.1
2011	4159	Southdown	156	1	R	*BB*	2.2	1.7	78.4	17.9	102	3.6	15	19.8	0.6	67
2011	4160	Merino	861/09	1	E	*AA*	2.36	2	84.8	18.9	115	3.9	16	20.5	0.6	43.8
2011	4162	Merino	742/08	1	E	*AA*	2.4	1.6	65.5	17.9	102	3.7	29	20.6	0.7	98.8
2011	4163	Merino	901/08	1	R	*AA*	3.12	2.4	78.3	18.5	85	4.3	19	23.5	1.7	74.7

2011	4164	Merino	833/09	1	R	*BB*	2.24	1.8	81.4	16.3	93	3.2	10	19.5	0.4	77.9
2011	4165	Merino	599	1	E	*BC*	2.68	1.8	66.7	19.8	98	4.6	26	23.5	3	103
2011	4169	Merino	763	2	E	*AA*	3.26	2.4	73.9	19.9	99	3.2	30	16.2	0.3	90.1
2011	4170	Merino	763	2	E	*BB*	3.58	2.8	78.1	18.7	119	4.2	33	22.3	1.5	64.1
2011	4172	Merino	886	2	E	*AA*	1.9	1.4	73.1	17.6	88	3.4	33	19.4	0.4	90.3
2011	4173	Merino	886	2	E	*AB*	2.46	1.9	75.5	18.2	94	3.6	34	19.8	0.55	73.4
2011	4174	Merino	799/08	2	R	*BB*	2.66	2.2	82.6	17	92	3.5	20	20.9	0.3	65.8
2011	4175	Merino	799/08	2	E	*AA*	1.96	1.5	80.4	16.1	95	3.1	12	19	0.05	62.8
2011	4177	Merino	828/09	1	R	*AA*	2.1	1.4	68.7	14.8	77	3.4	9	22.9	0.3	83.1
2011	4178	Merino	928/08	1	R	*BB*	2.16	1.5	70.1	15.8	79	4	27	25.6	0.5	75.6
2011	4188	Merino	925/08	2	E	*AB*	2.26	1.6	69.9	15.2	63	2.7	21	17.5		86.4

Appendix D

Copy of Publications Arising from this Study